ARI Research Note 96-59

The Effects of Performance, Individual Differences, and Arousal on Feedback-Seeking Behavior in a Novel Computer Based Task

Roger B. Rensvold
Georgia Institute of Technology

Research and Advanced Concepts Office
Michael Drillings, Acting Director

April 1996

United States Army
Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Research accomplished under contract
for the Department of the Army

Georgia Institute of Technology

Technical review by

Joseph Psotka

REPORT DOCUMENTATION PAGE

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED (from. . . to)
1996, April	Final	June 1992-October 1994

4. TITLE AND SUBTITLE	5a. CONTRACT OR GRANT NUMBER
The Effects of Performance, Individual Differences, and Arousal on Feedback-Seeking Behavior in a Novel Computer Based Task	MDA903-92-K-0107
	5b. PROGRAM ELEMENT NUMBER 0601102A
6. AUTHOR(S)	5c. PROJECT NUMBER B74F
Roger B. Rensvold (Georgia Institute of Technology)	5d. TASK NUMBER 2901
	5e. WORK UNIT NUMBER C10

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Georgia Institute of Technology School of Management 755 Ferst Drive Atlanta, GA 30332-5600	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. MONITOR ACRONYM
U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: PERI-BR 5001 Eisenhower Avenue Alexandria, VA 22333-5600	ARI
	11. MONITOR REPORT NUMBER Research Note 96-59

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

COR: Michael Drillings

14. ABSTRACT *(Maximum 200 words)*:

This report focuses on how situational characteristics affect feedback seeking. It also examines how individual differences in feedback propensities affect feedback seeking. Proposed antecedents of feedback eliciting (overt feedback seeking) were examined utilizing a computer-based technique that permitted objective measurement of the behavior. A 2 X 2 research design was used, with two levels of social presence (an observer present or absent) and two different task rationales (evaluating the task, evaluating the participant). Feedback eliciting was operationalized in two complementary ways--as the number of times the participant elicited feedback, and as the number of seconds he or she spent examining feedback information. Two types of feedback were examined: (1) outcome feedback, or information about level of performance, and (2) process feedback, or information about how to improve performance. Separate hypotheses were formulated for each. in addition, to social presence and task rationale, the following variables were hypothesized to affect feedback eliciting: performance, arousal (measured as state anxiety), external feedback propensity, task-specific internal feedback ability, task familiarity, internal feedback propensity, self-esteem, locus of control, tolerance for ambiguity, and need for achievement.

15. SUBJECT TERMS

Performance Feedback-seeking Computer-based task

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT	17. ABSTRACT	18. THIS PAGE			
Unclassified	Unclassified	Unclassified	Unlimited	90	

THE EFFECTS OF PERFORMANCE, INDIVIDUAL DIFFERENCES, AND AROUSAL ON FEEDBACK-SEEKING BEHAVIOR IN A NOVEL COMPUTER-BASED TASK

Roger B. Rensvold
Georgia Institute of Technology

Technical Report 94-2
October 1994

Prepared with the support of the United States Army Research Institute for the Behavioral and Social Sciences, Office of Basic Research, Contract MDA 903-92-K-0107. The views opinions, and or findings contained in this report are those of the author and should not be construed as an official Department of the Army position, policy, or decision.

This technical report is based on a doctoral dissertation prepared by the author under the direction of David M. Herold.

OVERVIEW

This report presents the results of doctoral thesis. Its major focus is on how situational characteristics affect feedback seeking. It also examines how individual differences in feedback propensities affect feedback seeking.

Proposed antecedents of feedback eliciting (overt feedback seeking) were examined utilizing a computer-based technique that permitted objective measurement of the behavior. A 2 x 2 research design was used, with two levels of social presence (an observer present or absent) and two different task rationales (evaluating the task, evaluating the participant). Feedback eliciting was operationalized in two complementary ways; as the number of times the participant elicited feedback, and as the number of seconds he or she spent examining feedback information. Two types of feedback were examined, outcome feedback, or information about level of performance, and process feedback, or information about how to improve performance. Separate hypotheses were formulated for each.

In addition to social presence and task rationale, the following variables were hypothesized to affect feedback eliciting: performance, arousal (measured as state anxiety), external feedback propensity, task-specific internal feedback ability, task familiarity, internal feedback propensity, self-esteem, locus of control, tolerance for ambiguity, and need for achievement. Multiple regression analysis showed that only performance, need for achievement, and state anxiety were significant as predictors of feedback eliciting. Tolerance for ambiguity, internal propensity, task familiarity, and social presence were components of significant two-way interactions that accounted for additional variance in feedback eliciting. An exploratory investigation identified age as an additional predictor of feedback eliciting, plus interactions involving sex, age, and work experience.

ORGANIZATIONAL FEEDBACK: LITERATURE REVIEW AND RESEARCH HYPOTHESES

Feedback-Seeking in Organizations

The importance of performance feedback in organizations has been recognized ever since the earliest days of "scientific management." F.W. Taylor (1916) outlined a program of systematic observation, task analysis, personnel selection, training, individual performance documentation, and remedial training (feedback), all designed to optimize the efficiency of industrial workers. McGregor (1960) characterized the assumptions and methodologies of Taylor and other early theorists as "Theory X" management; McGregor's own "Theory Y" also implicitly recognized the importance of feedback in its assertion that people have an innate desire to be productive. Information about *how* productive one has been; i.e., feedback; is essential to the satisfaction of that desire. Comprehensive reviews by Ammons (1956) and Chapanis (1964) present evidence that feedback dependably enhances both performance and motivation.

For many years feedback was viewed as simple "knowledge of results" (Annett, 1969) and most research was confined to the laboratory. Recent research in both labs and actual organizations has expanded and elaborated the construct considerably. Greller and Herold (1975) identified five sources of feedback in organizations (organization, supervisor, coworkers, task and self), and established a framework for understanding individual feedback preferences. Jablin (1979) characterized feedback content in terms of its perceived sign, accuracy, quantity, timing and specificity, and whether it addressed work processes, worker behaviors, or outcomes. In a similar vein, Wroten (1980) characterized feedback content as consisting of comparative, evaluative, or prescriptive information. Ilgen, Fisher and Taylor (1979) proposed a model of the effects of feedback on recipients, and later provided an important theoretical linkage between organizational feedback and control theory (Taylor et al, 1984.) More recently, Fedor (1991) extended the model by incorporating elements of Fishbein and Ajzen's (1975) theory of reasoned action and Petty & Cacioppo's (1986) theory of persuasive argumentation.

In addition to feedback sources and recipient responses, research has examined feedback mechanisms; that is, how recipients acquire knowledge about their performance. In general, feedback can be either "sent" or "sought." Sent feedback is information provided by a source that was not actively elicited by the recipient. It may be provided by

the task itself, as when an inattentive driver hits a curb and receives both tactile and auditory feedback. It may be provided by a supervisor as part of an established management procedure, or by a teacher following a course of instruction. Sought feedback, on the other hand, is actively acquired by the recipient in response to a felt "need to know" (Ashford & Cummings, 1985). This process of active acquisition, usually called feedback-seeking, was the dependent variable of interest in the research reported here.

Feedback-seeking by the prospective recipient can follow either a direct or an indirect strategy (Ashford & Cummings, 1983). The direct strategy, generally referred to as feedback eliciting, involves some overt behavior intended to acquire feedback from a source, such as asking a supervisor for information about one's performance. The eliciting behavior is readily identifiable by the source, as well as by observers, as an attempt to obtain performance-related information. The indirect strategy, on the other hand, involves behavior that is less readily identifiable as feedback-seeking, such as paying particular attention to a supervisor's tone of voice, repositioning one's workstation to "eavesdrop" on coworkers, attending social functions where work is discussed, etc. The various manifestations of the indirect strategy are generally referred to as feedback monitoring.

Feedback eliciting is important as a way of acquiring information, and also as a technique for impression management. Larson (1989) suggested that eliciting may be one way that recipients "defuse" anticipated negative feedback by getting it out in the open before the source "explodes." Ashford and Tsui (1991) demonstrated that managers who actively elicited negative feedback not only had a more accurate understanding of how their work was evaluated by the sources, but also were evaluated more highly by the sources. One good way to gain an appreciation for the complexity of feedback is to contemplate this impression-management function of elicitation. Not only is the recipient's performance a factor, but so is his or her evaluation of it; as well as assumptions or uncertainties about the source's evaluations, the possible consequences of negative feedback, etc. (Fedor, 1991).

Ashford (1986) investigated performance feedback from a resource perspective, arguing that the perceived need for feedback, as predicted by a number of personal and organizational variables, would in turn predict both eliciting and monitoring. Although Ashford's model stimulated a flurry of research, her hypotheses were not particularly well supported in her own published work. Only five of 20 correlations between antecedents and behaviors were significant in the direction hypothesized; this despite the fact that the hypotheses were both well grounded in theory and intuitively reasonable. One possible

reason for these results was *reliance upon self-reports of feedback-seeking behavior.* Although an investigator frequently has no other way of getting data, self-reports are particularly problematical in feedback research, for the following reasons.

First of all, self-reports of behavior are necessarily retrospective. Although a survey question may be phrased in the present tense, the subject is not really being asked what he or she *is* doing (the investigator already knows that - the subject is filling out a questionnaire) but rather what he or she *typically* or *usually* does. The answer to such a question can only be based on what the subject *has* done. Retrospection is subject to a multitude of biases that have received close attention from researchers in the area of performance appraisal (Feldman, 1981: Ilgen & Feldman, 1983). One source of bias is accessibility. If a person has recently gone to the boss for feedback, he or she may retrospectively judge the frequency of that behavior to be greater than it actually was. Another source of bias is salience. An especially vivid episode (e.g., a request for feedback that produces extremely unwelcome news) may bias the reported frequency of feedback-seeking in the upward direction.

In addition, it is possible that some subjects have difficulty differentiating between feedback-seeking and other information acquisition. Consider one item, part of a scale used to assess feedback monitoring (Fedor, Rensvold & Adams, 1992): "I keep my ears open in case the (source) has any additional information." Suppose a supervisor is discussing general job requirements with a group of workers that includes the respondent. Although the supervisor may have no intention of conveying information about individual performance, it may be interpreted as such by the respondent, especially if he/she is feeling apprehensive about performance. The process of mentally comparing objective job information with an individual self-assessment may give impersonal job information the cognitive flavor of sent feedback.

Self-reports of feedback-seeking behavior may also be biased by the so-called *consistency motif*, or by *social desirability* (Podsakoff and Organ, 1986.) That is, a respondent may be motivated to report behavior that is consistent with his or her previously expressed attitudes, or that will be perceived by others as appropriate or desirable. Fedor et al (1992), for example, found that an individual propensity toward feedback (external propensity; see below) predicted subsequent self-reports of feedback-seeking behavior. There are three possible mechanisms that may explain this result. First, this may have been the manifestation, in emitted behavior, of a stable individual difference, totally independent of situational factors and measurement processes. Second, the respondents, having

reported their propensities, may have been motivated to report consistent behaviors. Third, social norms may have influenced reports of both feedback propensities and feedback-seeking behaviors. (The organization was a military flight school, which emphasized self-reliance, moral courage, and individual decision making.) Of course, some combination of all three mechanisms may have been at work.

Problems also arise when asking a feedback *source* to recount instances of feedback-seeking by a ***recipient***. As noted above, the cognitive processes involved are the same as those involved in performance appraisal (Ilgen and Feldman, 1983). The feedback source, like the recipient, may selectively remember particularly salient instances, more recent instances, or misinterpret information-gathering as feedback-seeking. If the source is required to provide exceptionally close, detailed supervision, as in a training program, he or she may be unable to distinguish between "sent" feedback provided as an instructor, and feedback elicited by the source: for example, Fedor et al (1992) found no correlation between the frequency of feedback-seeking reported by student pilots, and the same behavior observed from the perspective of their instructors.

Because of the theoretical and practical problems inherent in self-reports of behavior, it seemed appropriate to utilize external observations of behavior whenever possible. The research proposed here took that approach. The dependent variable of interest was limited to feedback eliciting, an emitted behavior that could be observed and objectively measured by a person other than the feedback-seeker.

Outcome and Process Feedback-Seeking

Hackman and Oldham's (1976) influential "core dimensions model", as operationalized by their Job Diagnostic Survey (JDS), stimulated research on the relationships between job characteristics and various job outcomes, notably performance and satisfaction. Feedback was conceptualized as one of the five "core" dimensions characterizing a job. When numerous studies yielded results inconsistent with the theory, Wroten (1980) proposed that the fault lay with the feedback construct, which he asserted was multidimensional, rather than unidimensional as Hackman and Oldham had hypothesized. His research supported this notion.

Wroten (1980) proposed three functions for feedback: comparison, evaluation, and prescription. Feedback is comparative to the extent that it is based on some standard, such as an objective goal, or the performance of others. It is evaluative to the extent that it is either positive or negative. For example, when a foreman tells a worker "You're ten

minutes late" the statement is, in itself, only comparative feedback. Only when the additional message "And that's bad" is added, either implicitly or explicitly, does the feedback become evaluative. It is more evaluative to the extent that the foreman expresses the exact degree of his displeasure.

Comparison and evaluation pertain to outcome, or what the feedback recipient actually did. The prescriptive dimension pertains to process, or what the recipient should do. Prescriptive feedback has the function of coaching or instructing. Comparison and evaluation indicate that performance needs to be changed; prescription indicates how to change it.

Comparison and evaluation seem more closely related to each other than either one of them is to prescription. Although one could make a case for examining comparison and evaluation separately, the conceptual difference between them seems less important than the difference between either of them and the construct of prescription. Therefore, for the purposes of this study, feedback having both comparative and evaluative functions was combined under the heading of *outcome* feedback. Feedback having only a prescriptive function is hereafter referred to as *process* feedback.

The feedback literature has consistently made a clear distinction between outcome and process feedback, although the former has received more attention (e.g., Ilgen and Moore, 1982). The two types have been shown to have different effects on performance-related outcomes. Early et al (1990) found that process feedback (through an interaction with goal setting) had a greater effect on the formulation of task strategy than outcome feedback, while outcome feedback had a greater effect on effort. Although conceptually distinct, both types of feedback seem to be required for optimum performance. Kim (1984) found that the concurrent presentation of outcome and process feedback resulted in the greatest increase in performance, while Battman (1988) attributed the lack of process feedback in his experiment to the limited effectiveness of outcome feedback in improving performance.

Nomenclature

We have differentiated between two types of feedback-seeking, monitoring and eliciting, and stated that the research reported here focused on eliciting. Further, we have differentiated between outcome feedback and process feedback. The dependent variables of interest below were (1) the frequency with which outcome feedback was elicited (outcome feedback eliciting), (2) the frequency with which process feedback was elicited

(process feedback eliciting), and (3) the relative proportion of the two frequencies with respect to total feedback eliciting. The discussion below is facilitated by use of the following abbreviations[1]:

OFE = outcome feedback eliciting

PFE = process feedback eliciting

TFE = total feedback eliciting = (OFE + PFE)

OER = outcome eliciting ratio = (OFE / TFE)

Any reference to feedback-seeking, without qualification, should be understood as including both monitoring and eliciting as processes, and both outcome and process information as objects.

Performance as a Predictor of Feedback Eliciting

At least two studies support the idea that higher levels of performance result in less feedback-seeking. Ashford (1986) found that negative beliefs about goal attainment predicted higher (self-reported) levels of both eliciting and monitoring; that is, respondents who characterized their performance as inadequate were more inclined to seek feedback than those who felt their performance was adequate. This finding contradicted Ashford's hypothesis that those who held negative beliefs about goal attainment would avoid feedback out of fear of receiving negative evaluations. In Ashford's study, the results indicated that the value of feedback, as a resource that the subjects needed to improve performance, was generally high enough to justify the possible costs (embarrassment, etc.) of seeking it.

Ashford (1986) used the recipient's own impressions of his or her performance as the predictor of feedback-seeking. Fedor et al (1992) used an objective measure, namely the grades given to flight students by their instructors. In one of the two phases of training studied, this grade was a significant negative predictor of feedback eliciting. Those who flew better reportedly asked for feedback less often. It should be noted that these data, like Ashford's, were retrospective self-reports of behavior, subject to the biases discussed above.

In addition to information gathering, the act of eliciting sometimes serves an additional function; namely, impression management. Feedback sources, wishing to avoid unpleasant confrontations, are generally reluctant to impart negative performance information (Larson, 1989). As a result, dissatisfaction with poorly-performing

[1]See Appendix for expansions of all abbreviations.

subordinates tends to build up until superiors feel compelled to take action. The action taken is usually an angry outburst of adverse feedback replete with personal attributions. Recipients are aware of this repression/explosion dynamic and may attempt to "defuse" it by seeking feedback when they are aware that their performance has been below par. This impression-management function of feedback-seeking also suggests that those performing at lower levels should seek more feedback. Taken together, all the arguments above support the following hypothesis:

H1: Performance will be negatively related to feedback eliciting.

The notion behind Ashford's (1986) unsupported hypothesis - that the anticipated cost of hearing negative feedback may outweigh the value of the feedback - suggests that performance level may affect the relative frequencies of process feedback eliciting and outcome feedback eliciting. PFE may be seen as incurring fewer personal costs, and as being "safer," than OFE, for the following reasons.

Feedback sources are reluctant to give negative feedback, and recipients are aware of this reluctance (Larson, 1989). Yet recipients may desire feedback, for either performance-improvement or impression-management reasons (or both). Asking "How could I do this job better?" (PFE) allows for the possibility of less personal, less evaluative feedback than asking, "Have I been performing up to standards?" (OFE). PFE alleviates reluctance on the part of the source by making it possible to discuss the task without first having to make blunt, comparative, and possibly negative statements about the recipient's performance, as would be required by OFE.

Process feedback from external sources is more useful to a poor performer than outcome feedback, because the latter is easier for the person to generate for him- or herself. Outcome feedback is generated by comparing goals and outcomes. A large discrepancy between goals and outcomes is easier to detect than a small discrepancy, and produces a quicker, more accurate judgment about performance. Therefore, the poor performer has less need to seek outcome feedback from external sources - he or she already has a good idea what it would be. Instead, the focal person is more inclined to seek information that would be immediately useful in improving performance; i.e., process feedback. The relative usefulness of outcome and process feedback for the poor performer predicts a higher frequency of PFE than OFE.

There is also a difference between the PFE and OFE that relates to possible perceptions of the recipient's motivation. OFE (e.g., "Have I met standards?") does not

necessarily imply an interest in improvement; the question even may be a prelude to quitting. Other cues, such as appropriate posture and tone of voice, are necessary if the elicitor wishes to convey a desire to better his / her performance. On the other hand, PFE (e.g., "What changes should I make in the way I work?") does imply a commitment to improvement, and tacitly invites the source to be sympathetic and supportive.

In summary, poor performers may be motivated to seek feedback for purposes of impression management. They will simultaneously attempt to create the most favorable impression, obtain the most valuable information, and minimize the risk of hearing something that may be damaging to their self-image. This leads poor performers to elicit process feedback in preference to outcome feedback. On the other hand, good performers have little need for process feedback, since they obviously know the correct procedures. They may, however, desire outcome feedback as a source of recognition and reinforcement; that is, they want to be told how good they are. Together, these considerations lead to the following hypotheses:

H2: As performance decreases ,the ratio of outcome to process feedback eliciting increases. (Performance is negatively related to OER)

This relationship is expected to be strongest when another person is aware of the feedback eliciting, and when the recipient believes that the other person's impression of them is important. On the other hand, even when the recipient is alone, he or she is still a witness to his or her own performance. When task performance is highly salient, and one suspects that his or her performance is poor, one may still be inclined to seek process feedback instead of outcome feedback. This strategy may offer recipients the prospect of improving performance, while avoiding the anticipated bad news about his or her performance so far.

Individual Differences

Feedback-specific individual differences

It is desirable to identify individual differences that are specific to feedback, and thus closer to the behaviors of interest than are more general constructs such as tolerance for ambiguity, etc. This would permit investigators to describe individuals in terms of their attitudes toward feedback, and to do so quickly and parsimoniously. Herold and Parsons

(1985) have identified three individual differences pertaining to feedback, two propensities and one ability: internal feedback propensity, internal feedback ability, and external feedback propensity. Internal feedback propensity (referred to as internal propensity) is a measure of the individual's predilection for self-generated feedback, or the desire to "figure it out for myself." It may also reflect self-reliance and a general lack of confidence in the opinions of others, especially opinions regarding one's work performance. Apart from a preference for self-generated feedback, there exists the question of how much confidence one has in it; that is, having "figured it out for myself," how confident am I that my conclusions are correct? Internal feedback ability (referred to as internal ability) is the measure of this confidence. Finally, external feedback propensity (referred to as external propensity) reflects a preference to get feedback from sources external to one's self, such as superiors and coworkers.

The wording of the items constituting the external propensity scale (e.g., "I like getting frequent feedback from others concerning my performance." Herold & Parsons, 1985) leads to a clear expectation about behavior: those who score high on the scale should be more disposed to engage in feedback eliciting, or at least enjoy feedback more when they receive it. In a longitudinal study of military student pilots, Fedor et al (1992) found exactly this relationship; students who reported higher external propensity, also reported engaging in more feedback-seeking behavior. The external propensity scale specifically addresses respondents' attitudes toward getting feedback from other people, and it may be that the measured attitude is strongest when the feedback source is human. However, it is reasonable to think that the propensity may generalize to other sources as well; that is, a man who is predisposed to ask other people for feedback may also be more inclined to read instruction manuals, compare his sales performance with industry-wide statistics, and peek at the answers to crossword puzzles.

> *H3: External propensity will be positively related to total feedback eliciting (TFE).*

To date, neither theory nor empirical results offer any rationale for hypothesizing a differential effect of external propensity on OFE versus PFE.

Internal ability is a measure of confidence in one's ability to generate one's own feedback; e.g., to know "what I am trying to do and how well I am proceeding towards my goal." (Herold and Parsons, 1985). Of the three propensities, internal ability has always been the most problematical. This construct, which reflects one's confidence in

one's own ability to generate accurate feedback, has consistently had the lowest reliability of the three, and the greatest variance in factor loadings across samples (Herold et al, 1994). Moreover, internal ability seems to be task-specific. Research by Herold, Parsons and Rensvold (1991) suggests the internal ability may develop over time in a training environment, supporting the reasonable assumption that internal ability is higher with respect to familiar than to unfamiliar tasks. This task-specific aspect of the construct seems most relevant to feedback-seeking behavior, and all references here to internal ability should be understood as task-specific.

H4a: Familiarity with a task will be positively related to internal ability.

Since eliciting feedback from another person entails a risk of embarrassment, self-generating one's own feedback is a lower-risk alternative to asking. But in order to be useful, feedback must be accurate. Therefore, the perceived utility of substituting self-generated feedback for feedback from others depends upon the focal person's estimate of the accuracy of the former compared with the latter. High internal ability indicates that the person ascribes high accuracy to self-generated feedback; therefore, *ceteris paribus*, those with high internal ability should be more inclined to rely on their own estimates of their performance, and be less inclined to ask others. (This argument assumes that the primary motivation for feedback-eliciting is to obtain information; the presence of other motives, such as need for praise or recognition, may weaken the hypothesized relationship.)

H4b: Internal ability will be negatively related to total feedback eliciting.

Combining H4A and H4b produces the following:

H4c: Task familiarity will be negatively related to total feedback eliciting.

Internal propensity appears at first glance to be the opposite of external propensity. A representative item from the scale states, "What I think of myself and my work is more important to me than what others think." (Herold and Parsons, 1985). As was argued above, the high external likes feedback from others. In contrast, the high internal likes to generate his or her own feedback. Again neglecting the distinction between OFE and PFE, the hypothesis for internal propensity as a main effect seems to be straightforward:

H5a: Internal propensity will be negatively related to total feedback eliciting.

Yet this hypothesis may be too simplistic. Internal and external propensity are not the endpoints of a single continuum; rather, factor analysis has consistently demonstrated that the two constructs are nearly orthogonal (Herold et al, 1994). So while it may be tempting to argue the converse of H3 for internal propensity, it may also be wrong.

The ability to generate one's own feedback is contingent upon knowing enough about the task to do so. The information contained in feedback is a resource (Ashford, 1986), and a person needs a certain amount of this resource before he or she can self-generate feedback. Therefore, in the case of completely unfamiliar tasks, i.e., tasks associated with low levels of internal ability, high internals should elicit more feedback than low internals. Such people want to be judges of their own performance; in order to achieve this capability, they want to learn as much about the task as possible, as quickly as possible. The pilot study reported earlier (Herold et al, 1991) found exactly this effect for a novel task. The effect may reverse as people acquire task experience. As high internals acquire the expertise they need to generate their own feedback, they may be governed by the main effect for internal propensity and engage in less eliciting than low internals.

H5b. Internal ability will moderate internal propensity as a predictor of feedback eliciting. When internal ability is low, higher levels of internal propensity will result in more TFE. When internal ability is high, higher levels of internal propensity will result in less TFE.

Other Individual Differences and Feedback Eliciting

A large number of individual differences, all of them predating the feedback-specific differences discussed above, have been investigated as predictors of behavior. Four seem particularly relevant to feedback-seeking; self-esteem, locus of control, need for achievement, and tolerance for ambiguity. All four are personality constructs; i.e., they relate to cognitive styles and behavioral predispositions that are usually considered to be independent of context. They are included in the present study for two reasons. First, it is desired to investigate their relationships to feedback-seeking behavior using the research methodology proposed here. Much previous research has relied on retrospective self-

reports of feedback-seeking, subject to the biases discussed above; the method to be used in this study permits objective measurement. Second, it is desired to compare and contrast the predictive validities of the personality constructs with those of the feedback-specific constructs, and to examine the correlations among the constructs.

Self-Esteem (SE) is defined by Brockner (1988) as "...(a) trait referring to individuals' degree of liking or disliking of themselves. Thus, the essence of self-esteem is the favorability of individuals' characteristic self-evaluations." (pg. 11) A number of studies have demonstrated links between SE and both task performance and task-related behaviors, including feedback-seeking.

SE is hypothesized to have two effects on task performance, one due to effort and the other to information searching. Because high-SE persons desire to perform at a level consistent with their self-image, they tend to work harder and perform better on a variety of tasks (Dipboye, 1977; Korman, 1970). But performance depends upon more than effort; to a degree dictated by the nature of the task, it is also dependent upon the performer's access to relevant information, including feedback. Lack of information sometimes impairs the performance of high-SEs. Because of their greater self-confidence, high-SE people seek less information before attempting a task, and less feedback while performing it.

Weiss and Knight (1980) presented subjects with a numerical problem-solving task, with the option of seeking additional information before proposing a solution. High-SEs sought less information, and did more poorly, than low-SEs. In an extension of this research, Knight and Nadel (1986) presented subjects with a computerized business-policy task. Subjects chose investment policies, selected the frequency with which they would receive feedback about their decision, and were given the opportunity to change their policies in response to the feedback. High-SEs elected to receive feedback less often, and stayed with "losing" policies longer before changing them.

Both Weiss and Knight (1980) and Knight and Nadel (1986) conducted laboratory studies. A longitudinal field study (Fedor, Rensvold, and Adams, 1992) replicated the observed relationship between SE and feedback-seeking. In a sample of 137 student pilots, SE was negatively associated with self-reports of asking instructors for feedback. Based on the evidence, it is hypothesized that:

H6a: Self-esteem will be negatively related to total feedback eliciting.

Brockner (1988) suggests that high-SEs have more to lose as a consequence of receiving feedback than low-SEs, due to the greater discrepancies between negative

feedback and their self-images. This is supported by research showing that high-SEs are less accepting of criticism than low-SEs (Sweeney and Wells, 1990: Campbell, 1990), which may be interpreted as indicating that high-SEs pay a higher price when confronted with disconfirming information about themselves than do their low-SE counterparts. This suggests that SE may have a differential effect on OFE and PFE. Asking for process feedback - i.e., how to improve task execution - does not necessarily invite the source's attention to the recipient's outcomes with respect to goals. Nor does it necessarily imply that the recipient is worried about the quality of his or her performance, since process-oriented questions may indicate nothing more than an interest in alternative procedures. Therefore, high-SEs may perceive PFE as less threatening than OFE.

H6b: High self-esteem individuals will request proportionately more process feedback, and less outcome feedback, than low self-esteem individuals. (SE will be negatively related to OER.)

Locus of Control (LOC) is an individual difference describing the degree to which a person believes that behavioral reinforcements are contingent upon his or her own behaviors or attributes (Rotter, 1966). The construct is bipolar, ranging from external to internal. Those at the external end of the continuum ("externals") attribute the events in their lives to outside forces, such as luck, fate, or the influence of powerful others. "Internals," on the other hand, believe that the things that happen to them are contingent on their own (internal) characteristics and behaviors.

As discussed above, one's attitude toward feedback-seeking depends in part upon the utility ascribed to the anticipated outcome; i.e., the usefulness of the information that one expects to gain. The attitude will also depend upon the anticipated utility of feedback-seeking behavior as an impression-management mechanism (Larson, 1989). Both judgments are affected by the degree to which the focal individual sees him- or herself as able to influence outcomes; that is, able to perform better if given feedback, or able to favorably influence a source through feedback-seeking. If performance quality and the opinions of others are perceived as being due to exogenous factors not under the individual's control, then he or she should place little value on feedback. Therefore externals should seek less feedback than internals.

This notion is supported by Feather and Volkmer's (1988) finding that externals preferred classroom situations characterized by little feedback; if feedback is perceived as irrelevant noise, then less of it is preferable to more. Other studies indicate that externals

sample less performance-relevant information than internals (Boone, Brabander and Gertis, 1991) and make less effective use of it to improve performance (DuCette and Wold, 1973).

H7a: An internal locus of control (high value of LOC) predicts higher levels of feedback eliciting than an external locus of control. (LOC is positively related to TFE.)

One's location on the internal-external dimension should differentially affect OFE and PFE. Externals believe that the success or failure of their efforts is attributable to outside forces. Although they may attempt to ascertain the nature of those forces, externals should have little interest in process feedback as such. Since the outcomes they experience are thought to bear little relation to their efforts, externals should have little interest finding out how to work more effectively. Internals hold the opposite opinion. On the internal-external dimension, the attitude toward process feedback should range from interest to indifference. The attitude toward outcome feedback, on the other hand, should range from interest to aversion. Internals may not enjoy receiving negative feedback any more than externals, but internals may be able to take refuge in the rationalization that the feedback will help them improve their performance. Externals have no such comfort. Since they believe they have little or no control over their performance, externals may tend to experience negative feedback as "pain without gain." Feather and Volkmer (1988) found some evidence supporting this notion; in their study of 80 students, externality was positively correlated with text anxiety, which presumably involves an aversion to being graded (that is, receiving outcome feedback).

H7b: Individuals with an external locus of control (low value of LOC) will request proportionately more process feedback, and less outcome feedback, than those with an internal locus of control. (LOC is negatively related to OER.)

Tolerance for ambiguity (TA) is defined as "the tendency to perceive ambiguous situations as desirable." (Budner, 1962: p. 29) By this definition, low-TA individuals should find ambiguous situations noxious and attempt to make them less so by acquiring additional information. It follows immediately that TA should be inversely related to feedback-seeking, and this has been empirically demonstrated (Ashford and Cummings, 1985; Bennett, Herold and Ashford, 1990; Fedor et al, 1992).

Ashford and Cummings (1985) found a negative zero-order correlation between feedback-seeking and TA ($p < 0.001$), and significant interactions between TA and the two independent variables of role ambiguity and contingency uncertainty. The interactions indicated that high TA reduced the magnitude of the positive relationships between the two variables and feedback-seeking. The same study indicated that the items used for TA loaded on two factors, identified as problem-solving TA and job-related TA. A reanalysis of the data utilizing these two types of TA as predictors (Bennett et al, 1990) demonstrated that job-related TA was negatively related to feedback-seeking about both performance and promotion potential. The effect was observed when both coworkers and supervisors were sources.

Fedor et al (1992), working with student pilots, found that TA had negative zero-order correlations with feedback monitoring and eliciting across two separate phases of flight training. The regression of TA on feedback eliciting in both phases produced weights that were negative but nonsignificant.

These results, although somewhat sketchy, are in agreement with one's intuitions about the meaning of TA and its effects on feedback-seeking. None of the studies cited differentiated between OFE and PFE, and there are neither theoretical nor intuitive bases for hypothesizing a differential effect of TA on the two; therefore, they are combined as TFE in the following hypothesis.

H8: Higher tolerance for ambiguity predicts lower levels of feedback eliciting. (TA is negatively related to TFE.)

Need for Achievement (nAch) has one of the largest literatures of all personality constructs, having been investigated with respect to motivation, economics, psychotherapy, women's liberation, social class differences, racial inequity, and education (Weiner, 1978). NAch is manifested as "(the desire) to accomplish something difficult. To master, manipulate or organize physical objects, human beings, or ideas. To do this as rapidly and independently as possible. To overcome obstacles and attain a high standard..." (Murray, 1938; quoted in Weiner, 1978). It has been shown that high-nAch persons set higher goals for themselves, work harder toward more specific goals, generally perform better, and demonstrate improved performance in response to feedback (Steers, 1975; Matsui, Okada, and Kakuyama, 1982).

At first glance, the relationship between feedback-seeking and nAch seems obvious; if feedback information is necessary to improve performance, then high-nAch

persons should engage in more feedback-seeking than low-nAch persons. At second glance, a portion of the quote from Murray raises doubts. If doing something "independently" is part of achievement, then perhaps evaluating and correcting one's own performance - i.e., self-generating feedback - is preferable to seeking feedback from others. The conjecture would benefit from additional support.

The support comes from extensive studies concerning task choice. It is well established that high-nAchs prefer tasks of intermediate difficulty, while low-nAchs prefer tasks that are either comparatively easy or comparatively difficult (Weiner, 1978). There are two explanations. The first is based on an algebraic formulation involving motivation to succeed, motivation to avoid failure, and the perceived probabilities of success and failure (Atkinson, 1957, 1964). If the assumptions of the model are granted, the observed behavioral dispositions follow automatically.

The second explanation, which is more useful to the present discussion, involves task diagnosticity (Weiner, Heckhausen, Meyer, and Cook, 1972). A task is diagnostic to the extent that it permits ability to be inferred from performance. If a task is low in diagnositicity, then ability will have little effect on outcomes; an example is playing a lottery. If a task is high in diagnosticity, then ability will have a large effect on outcomes; an example is playing golf. Difficult and easy tasks are low in diagnosticity, because most people perform poorly on the former, and well on the latter. Tasks of intermediate difficulty, on the other hand, are high in diagnosticity, because they exhibit a range of outcomes that yield information about individual abilities (Meyer, Folkes and Weiner, 1976). According to Weiner (1978), this explains why high-nAchs prefer tasks of intermediate difficulty; the tasks provide reliable information about their personal competencies. Such information is a valuable resource. Using it, the focal person can decide upon a course of action, such as whether to train, study, or avoid similar tasks in the future.

Weiner's (1978) theory suggests that high-nAchs have more desire to obtain performance-related information than low-nAchs, and this explains their choice of tasks. In other words, high-nAchs set themselves up to obtain feedback whenever possible. Even in situations where the task is assigned rather than being freely chosen, it is reasonable to expect that high-nAchs will be more proactive in their search for performance-related information than low-nAchs. Not only do they want to maximize their present performance (the original conjecture), but they also want to learn more about their abilities, and thus infer something about their future performance.

H9a: Need for achievement will be positively related to total feedback eliciting.

The quotation from Murray (1938) concerning the various manifestations of nAch suggests that the ability to act independently is a part of achievement. It also suggests that the level of nAch may have an effect on one's choice of whether to seek process or outcome feedback. A high-nAch is hypothesized to be more concerned with performance, and therefore expected to elicit more outcome feedback, than a low-nAch counterpart. On the other hand, a high-nAch may also be more concerned with freedom of action. Developing his or her own approach to a problem, and carrying it out independent of coaching, may be an important component of the satisfaction he or she derives from the task. If this is true, then one would expect to find proportionately less process feedback eliciting. A high-nAch wants to do well, but in addition he or she wants "to do it my way." This argument leads to the following speculative hypothesis.

H9b: Higher need for achievement predicts a proportionately higher frequency of eliciting outcome feedback versus process feedback. (NAch is positively related to OER.)

Situational Factors Affecting Feedback Eliciting

So far, the discussion has centered on individual factors hypothesized to affect feedback-seeking behavior. Tolerance for ambiguity, need for achievement, and the other constructs discussed in the preceding section are all individual differences having some standing in the psychological literature. Even task performance has individual components consisting of experience, aptitude, psychomotor skills, and related factors.

We now wish to consider factors that are specific to the task and task situation. Many could be listed, including time constraints, contingent rewards, specific characteristics of the task, and feedback availability. Availability will be neglected here in favor of factors that affect the individual's decision to seek feedback under conditions when he or she believes it can actually be obtained; that is, with availability as a precondition. The effects of believing that feedback is available may, however, be moderated by the salience of that belief. If asked, a person may assert that he or she can indeed obtain feedback in conjunction with a task, but later may lose sight of that fact when actually performing the task. This point is considered in more detail following hypothesis 10b, where certain theoretical caveats are addressed.

Differentiating situational factors affecting behavior from individual factors is not straightforward. Forces originating in the organization, group, or environment clearly must produce some change at the individual level before they can influence individual behavior. To use a simple physiological example, extreme temperatures do not have strict mechanistic relationships to performance, but rather produce distraction and annoyance (individual level responses) in addition to physical debilitation. In discussing situational factors affecting feedback eliciting, it is useful to focus on one individual-level response that covaries with a large number of situational factors, including threat, temperature, things producing apprehension or pleasure, etc. That response can be viewed as a highly generalized "function" that maps a variety of situational factors onto individual behavior. It is known as *arousal.*

Heightened arousal is associated with physiological changes such as increased pulse, respiration, sweating and skin conductivity, and psychological changes such as heightened perception and increased aggressiveness. Heightened arousal can be produced by a large number of factors such as exercise, crowding, heat, pain, and many other stimuli, but the discussion here will be limited to two. These two factors, which in turn have a large number of possible antecedents in organizations, are *social presence* and *evaluation apprehension.* In sum, the theoretical development presented here aims at a high degree of generalizability by using the following strategy. Many organizational factors affect social presence and evaluation apprehension; these, plus other factors, affect arousal; and level of arousal can be invoked as a proximal cause for changes in feedback-eliciting behavior.

The presence of other people produces a phenomenon known as social facilitation, the study of which established the discipline of social psychology (Triplett, 1898). A multitude of investigators have established that the presence of other people enhances the performance of well-learned tasks and inhibits the performance of unfamiliar tasks (for a review, see Geen, 1991). Although a number of distal mechanisms have been proposed, most authorities agree that the proximal cause of the social facilitation is an increase in general arousal, or drive level. Higher drive level increases the emission of well-habituated behaviors (dominant responses) relative to less-habituated behaviors. As a result, "...(T)he presence of others will cause increased emission of dominant responses, thus improving performance where such responses are correct (such as in the performance of familiar tasks), and impairing performance where such responses are incorrect (such as in the learning of novel tasks)." (Glaser, 1982; 266)

Although social facilitation makes unambiguous predictions about how task performance will be affected by the presence of others, it is less clear what predictions should be made concerning the feedback-seeking behavior arising from the same task. When one is immersed in a task, what constitutes the "dominant response"? It is suggested here that the dominant response is simply to continue doing the task, and not interrupting it to elicit feedback. Conversely, feedback eliciting is far from being a dominant response under any circumstances, since it is contingent upon judgments involving goals, uncertainty, risk, effort, and ego defense (Ashford & Cummings, 1983; Ashford, 1986) as well as impression management (Larson, 1989).

Because of the various costs involved (Rensvold et al, 1993) it is expected that feedback seeking will usually be a reasoned action, as opposed to an unconscious or automatic action. The necessary cognitive antecedent of a reasoned action is a conscious behavioral intention (Ajzen, 1989) based on attitudes, social norms, and perceived behavioral control. Formulating such an intention requires interrupting the person's preoccupation with the task; to that extent, the intention-formulating behavior is "novel." The "doing" of the task, even though it may have gone on for only a few seconds or minutes, is "habitual" in the sense that it will continue until interrupted. Therefore, it is anticipated that even if there is no *conscious* intention to avoid feedback, the mere presence of another person will reduce the frequency of feedback eliciting; further, this effect will be greater for novel tasks than for familiar ones. This effect was observed in a pilot study, in which the presence of an observer reduced the overall level of feedback eliciting, even though subjects were assured that the observer was not evaluating their performance; further, the feedback was elicited from a computer, not the observer. In summary, it is hypothesized that

> ***H10a: The presence of an observer inhibits feedback eliciting (is negatively related to TFE).***

Recapitulating, the causal relationships underlying H10a are as follows: social presence --(+)--> arousal --(-)--> feedback eliciting; that is, the presence of other people increases arousal, which in turn decreases feedback eliciting. Other factors are expected to increase arousal, notably evaluation apprehension.

Evaluation apprehension is known by various names, such as test anxiety and stage fright. It is a common human experience with an obvious implication for feedback-seeking behavior, even if the arousal mechanism is not specifically invoked. if one is

afraid of receiving an evaluation, one should be reluctant to ask for it. *Ceteris paribus*, high levels of evaluation apprehension should produce lower frequencies of feedback eliciting.

Evaluation apprehension is expected to covary with the construct of feedback-seeking cost, which is "the risk of embarrassment and loss of face inherent in inquiry " (Ashford, 1986: p. 471) due to the possibility of "hearing something that one would rather not know." (p. 470). Specifically, apprehension is expected to increase as the perceived costs of seeking feedback increase. Fedor, Mathieson and Adams (1990) demonstrated that the feedback-seeking costs associated with particular sources (i.e., supervisors) were negatively associated with intentions to ask those sources for feedback, and positively related to the intention to seek feedback from alternative sources. Fedor, Rensvold and Adams (1992) found that feedback-seeking cost was a significant negative predictor of self-reported feedback seeking. Neither of these studies measured evaluation apprehension directly, but it was very likely present as a mediator: that is, the perception of high costs -- e.g., a great risk of embarrassment and damage to one's self-image -- produced apprehension, which in turn inhibited the behavior.

Although the studies cited above involve evaluation by other people, the effects of evaluation are expected to extend even to situations where others are not present. People are competitive to a greater or lesser degree, and frequently test themselves via self-imposed tasks even when success or failure is both invisible and meaningless to others; witness the popularity of crossword puzzles. The theory that people need to be seen, and to see themselves, as competent is present, either implicitly or explicitly, in numerous models of motivation -- e.g. Maslow's (1943) esteem and self-actualization needs, Herzberg et al's (1959) achievement and recognition motivators -- and has great face validity[2]. The belief that one's task performance will result in some measure of one's competence is therefore expected to be arousing. Heightened arousal is, in turn, is expected to inhibit feedback eliciting via the arousal mechanism described in connection with H10a above.

H10b: The perception that one is being evaluated will be negatively related to total feedback eliciting.

---

[2]Which may explain the universal inclusion of these theories in textbooks, despite ambiguous experimental support and heavy criticism (Hall & Nougaim, 1968; House & Wigdor, 1967; and many others.)

In a manner analogous with H10a, the principle causal connections underlying H10b are as follows: evaluation --(+)--> arousal --(-)--> feedback eliciting; that is, the belief that one is being evaluated increases arousal, which in turn decreases feedback eliciting.

A few caveats must be entered. First, the demand characteristics of the task, or the outcomes contingent upon task performance, may be so salient that the opposite effects are obtained. For example, if an employee must master a task or be fired, it is reasonable to expect that the anxiety produced by that eventuality would result in more feedback eliciting, not less. Again, a person performing a task may interpret the presence of an observer as an invitation to engage in feedback eliciting, reasoning that if the other person was not intended to facilitate things in some way, he or she wouldn't be there. Even if the observer is not a source, the impulse may result in increased feedback-seeking from other sources. Given this interpretation on the part of the would-be recipient, observer presence may increase eliciting, not decrease it.

To consider yet another possibility: the task itself may be so involving that the person may lose sight of the fact that feedback is available. In this case, the high state of arousal produced by the task itself may overwhelm any effects attributable to observer presence, evaluation, or any other source. This effect would be explicable in terms of cognitive resources. The demands of the task would decrease the person's ability to self-evaluate and determine that feedback is needed, and/or to monitor the environment for feedback sources. (Kanfer & Ackerman (1989) invoked the model of limited cognitive resources to explain why subjects with established performance goals learned a complicated task more slowly than subjects without goals; self-monitoring of performance with respect to the goals tied up resources that otherwise would be allocated to learning the task.) In sum, the relationships set forth in the hypotheses above are expected to be observable only in the absence of extreme circumstances, high demand characteristics, assumptions concerning the normative value of feedback-seeking behavior, or extreme involvement with a highly demanding task.

It should also be noted that social presence and evaluation apprehension, despite their obvious relevance to organizational behavior, are not the only sources of arousal; further, arousal may have the same effects on behavior regardless of its source (Cottrell, 1972). For example, performing a task under time pressure may increase arousal, and decrease feedback eliciting. The same effect can be produced by threat. Arousal produced by threat increases the probability of well-rehearsed (i.e., dominant) responses *vis-a-vis*

novel responses, even when the former are totally inappropriate. This *threat-rigidity effect* has been observed in the behavior of individuals, groups and organizations (Staw, Sandelands and Dutton, 1981). The military provides an excellent example. Since earliest times, trainers have impressed upon recruits that "You'll fight the way you train." Under the threats and time pressures of combat, soldiers always emit their best-learned behaviors, even when it is obvious (to others) that to do so invites catastrophe (Hackworth, 1989). In terms of feedback eliciting, the obvious moral is that changes in the eliciting induced by heightened arousal are not necessarily due to either social presence or evaluation apprehension.

The hypotheses above posit relationships among social presence, evaluation, arousal, and total feedback eliciting (TFE). The theoretical foundation is not strong enough to support equivalent arguments for the ratio of outcome to total feedback eliciting (OER). There are, however, some grounds for expecting a relationship between evaluation and OER.

In the pilot study cited above, TFE was suppressed by the presence of an observer. In addition, OFE was suppressed more than PFE (means of 0.38 vs. 1.17). This differential effect is explicable in terms of an argument similar to the one given for the differential effects of self-esteem (above). Consider the possible consequences of asking for outcome feedback, as opposed to process feedback. The former behavior may draw attention to one's performance, or even bring to light performance information that was hitherto unavailable. In the pilot task, for example, performance information was not available to either the participant or the observer until the participant asked to see it. It follows that if one believes he or she is being evaluated, and is experiencing a degree of evaluation apprehension, then one will tend to avoid asking for outcome feedback. Asking for process feedback, on the other hand, does not have the effect of unearthing any hitherto unavailable performance information, nor does it necessarily communicate any concern about one's performance. Even people who perform at a very high level may be motivated to seek process feedback out of simple curiosity. A famous singer may, for example, be interested in hearing other singers' critiques of her technique, even if she has no intention of following their advice.

H11: Being evaluated on the basis of task performance decreases the frequency of outcome feedback seeking relative to process feedback seeking. (Evaluation is negatively related to OER.)

Interactions

The independent variables included in this research consist of three types of constructs. The first type consists of individual differences, such as self-esteem, need for achievement, and the feedback propensities. The second type was performance, which is a function of both individual differences, such as ability, and situational factors, such as task difficulty. The third type consists of situational factors, namely social presence and evaluation apprehension, that are expected to influence feedback-eliciting behavior through the arousal mechanism. Since one's assessment of a situation is the result of interactions between individual-level and situational factors (Bowers, 1973), it is reasonable to expect interactions among the variables presented here.

Although no specific hypotheses concerning interactions are proposed, there is ample reason to expect them, both within and between the different types. Within the individual differences, achievement need is hypothesized to be positively related to feedback eliciting, and self-esteem negatively related. Therefore, the effect of achievement need on feedback eliciting may be moderated by levels of self-esteem.

Within the situational factors, social presence and evaluation apprehension are hypothesized to be independent predictors of feedback eliciting; that is, the presence of an observer is thought to inhibit feedback eliciting, whether or not the would-be recipient believes he or she is being evaluated. In addition, being evaluated is thought to affect feedback eliciting, whether an observer is present or not. It is reasonable to expect that the effects of being evaluated would be more pronounced when one is being evaluated in the presence of an observer, which would imply an interaction between evaluation and social presence.

In addition, interactions may exist between performance and social presence. Lower levels of performance are hypothesized to increase feedback eliciting, due to the value of feedback in improving performance. The presence of an observer, however, may suppress this behavior through the arousal and ego-defense mechanisms discussed above; the effect would be an interaction between performance and social presence, characterized by a negative relationship between performance and feedback eliciting when recipients are alone, and a weaker or nonexistent relationship when observers are present.

METHODS

The experiment utilized data obtained from questionnaires, and data collected during the performance of a computer-based task. Three questionnaires were administered; one several days before the experimental task (the "survey"), one immediately before the experimental task (the "prequiz"), and one immediately afterward (the "postquiz"). The task was embedded in a 2x2 factorial design, with two levels of social presence (observed and not observed) and two evaluation rationale (evaluation of task, evaluation of the subject). Criteria were outcome feedback eliciting, process feedback eliciting, and the ratio of outcome to total feedback eliciting. Each criterion was operationalized in two different ways; as the number of separate instances of each behavior, and the number of seconds spent performing each behavior. The predictors were those discussed above; performance, external propensity, internal ability, internal propensity, self-esteem, locus of control, tolerance for ambiguity, need for achievement, and evaluation apprehension.

Participants

Participants were recruited from undergraduate management classes at the Georgia Institute of Technology, and received academic credit for their participation. Since both the survey questionnaire and the experimental task explicitly addressed feedback-seeking behaviors, steps were taken to prevent responses on the former from affecting performance on the latter. The two activities were scheduled for different days, and they were presented as separate experiments, with separate credit awarded for each. Participants were free to sign up for either one, or for both. This conservative approach resulted in the loss of experimental power, since 176 participants completed the survey, 154 the task, but only 129 completed both.

Demographics of participants completing both the survey and the task were as follows. Age ranged from 19 to 40, with a mean of 22.2 and a standard deviation of 2.52 years. Years of full-time work experience (self-reported) ranged from zero to 17, with a mean of 1.98 and a standard deviation of 2.68. Forty-two of the participants were female, 78 were male, and gender data were missing for two. Sixteen were sophomores, 60 juniors, 45 seniors, and one was a graduate student. Data on race were not collected.

Procedures

This section is an overview of the procedures utilized in this study. Details are deferred to the next section, where the survey, prequiz, laboratory task, task manipulations, and postquiz are discussed.

The participants were recruited in class, and completed two sign-up sheets; one for scheduled administrations of the survey, the other for laboratory sessions. The survey was administered to small groups in a classroom setting. The laboratory sessions were scheduled for one hour each, with one participant per session.

Internal propensity, external propensity, and internal ability were considered to be special, since they were all formulated as feedback-specific individual differences. In order to maximize the possibility of detecting differences in behavior attributable to these three variables, special care was taken to distribute them as evenly as possible among the four experimental manipulations (described below). The following procedure was used. The three variables were dichotomized using median splits, then crossed to produce a 2 x 2 x 2 classification scheme. Each participant was assigned to one of the eight classes based on his or her scores on the variables. Individuals within each class were then randomly assigned in equal numbers to each manipulation. If the number of individuals in a class was not evenly divisible by four, those in the remainder were treated as follows. A remainder of 3 were randomly assigned to manipulations I, IV, and either II or III, as determined by a coin flip. A remainder of 2 were randomly assigned to manipulations I and IV. A remainder of 1 was assigned to manipulation IV.

Upon arriving at the laboratory, all participants signed a statement indicating that they understood the nature of the task. Possible health hazards (repetitive motion injury and flicker vertigo) were explained and participants were given the opportunity to withdraw while still receiving full academic credit. Each participant then completed the pencil-and-paper prequiz.

After completing the prequiz, each participant viewed a videotaped presentation explaining the nature of the task. There were four presentations corresponding to the four experimental manipulations. Each contained deceptions concerning the purpose of the task, which were approved in advance by the institutional Human Subjects Committee. After viewing the presentation, the participant performed the experimental task on a Macintosh computer. After the computer program terminated, the participant completed the pencil-and-paper postquiz and was allowed to view a record of his or her task performance.

Finally, each subject was completely debriefed about the nature of the experiment, to include the deceptions used, and was pledged to secrecy.

Instrumentation

The Survey

Before performing the task in the laboratory, all participants completed a 141-item instrument known as the GTEC Student Survey. The survey included the following demographic information:

- Gender (coded female=1, male=2)
- Student identification number (Social Security account number)
- Year of birth (two-digit year)
- Cumulative years of full-time employment
- Date instrument was completed (month, day, two-digit year)

Other items on the survey assessed the constructs of external propensity, task-specific internal ability (based on previous experience with computer-based tasks), internal propensity, self-esteem, locus of control, tolerance for ambiguity, and need for achievement; i.e., the individual differences constituting the independent variables of the study. Other data were also collected; these were exploratory in nature, and not required to test the hypotheses presented above. Certain of these data are discussed at the appropriate points in the analyses and discussions that follow.

Physically, the survey consisted of reusable question booklets with a separate mark-sense answer sheets that participants filled out with pencil. The answer sheets were scored by machine.

The Laboratory Prequiz

Each participant was initially seated at a table in a corner of the laboratory and asked to complete the pre-experiment questionnaire, or prequiz. The table was equipped with a privacy partition and a box in which the completed prequiz was deposited. Physically, the prequiz consisted of a nonreusable paper form with multiple choice responses that participants marked with either pens or pencils. Data were transcribed from these forms to a data base by the experimenter.

Prequiz items assessed the participant's previous experience with computers, particularly the graphical user interface that was one of the defining characteristics of the

task. The participants were then given a description of the task that highlighted its particular demands, such as hand-eye coordination, but omitted specific details about how it was to be performed. Each then indicated how well he or she expected to do, yielding a measure of task-specific self-efficacy. This measure was distinct from the hypothesized predictor of task-specific internal ability discussed above, and was collected for exploratory purposes.

The Experimental Task

The experimental task was a variety of two-axis tracking exercise implemented on a Macintosh computer. Unlike other tracking tasks, the exercise required much more than simple hand-eye coordination. In order to satisfy the task criterion, each participant was required to examine all available information, access available feedback as necessary, and arrive at a systematic approach. The dependent variables of interest were participants' performance and feedback eliciting behaviors, which were recorded by the program.

The task had its origins in aviation training. All pilots must cultivate division of attention (DA); that is, the ability to check one flight parameter, make a control input if necessary, then check other parameters while waiting for the input to take effect. The opposite of division of attention is fixation. An aviator is fixating if she or he concentrates exclusively upon one discrepant parameter until it is corrected to his/her satisfaction, while ignoring others. One simple example will suffice. In straight-and-level flight, the three primary parameters are heading, altitude, and airspeed. A pilot must monitor all three parameters at all times. If the aircraft deviates from altitude (perhaps due to turbulence), a pilot with appropriate DA will adjust power or pitch attitude as necessary, then continue to scan all three parameters while the altitude deviation corrects itself. A fixating pilot, on the other hand, will make a control input, and then stare at the altimeter until the aircraft has returned to its assigned altitude. In the meantime, heading and airspeed may have exceeded their assigned values to an unacceptable degree.

In the early days of this research, it was conjectured that DA had two components; a skill that could be imparted by training, and an innate ability that placed an upper limit on DA, its acquisition, and its retention. Feedback-seeking propensities were thought to be related to the ability component of DA. One tentative hypothesis, for example, posited that those with high external propensity would be more receptive to performance-related task data, and thus would manifest better DA. Such relationships, if

demonstrated, would have immediate practical application to the selection of students for pilot training, since good DA is one of the attributes that every good aviator must possess.

Preliminary attempts to test these ideas using a laboratory-based helicopter flight simulator were abandoned after it was discovered that flying the simulator was much too difficult for subjects not having prior aviation training. Even intelligent and otherwise highly competent nonaviators could not keep it from "crashing." The investigator then began developing a simpler, microcomputer-based task that would require good DA to achieve a high score: the notion was that feedback propensities would predict DA, which in turn could be inferred from task performance. This line of inquiry was beset by theoretical and practical difficulties.

In time, the dependent variable of interest shifted away from measuring DA to measuring feedback-seeking behavior. The task was modified to give subjects the option of temporarily interrupting program execution, freezing the display, and requesting task-related information from a series of on-screen menus. These behaviors were recorded by the program as instances of feedback-seeking.

Figure 1 depicts the experimental display. The heavy rectangle represents the edges of the **window**, or area of the computer screen utilized by the program. **Crosshairs** located the center of the window. A small black arrow called the **cursor** was the conventional Macintosh symbol indicating the relative position of the computer mouse. A small circle, called the controller for reasons explained below, was controlled by the cursor. Two graphical objects, a **circle** and a **square**, moved up and down parallel to the left and right edges of the window. The object of the task was to keep the circle and square in contact with the **targets**, horizontal lines located midway up the window on the left and right sides.

Figure 1 About Here

Figure 2 depicts the relationship between cursor position and controller position. When the cursor was on the crosshairs the controller was superimposed on it. When the participant moved the cursor moved away from the crosshairs, the controller moved away by the same distance, at a 90 degree angle.

Figure 2 About Here

The circle and square were constrained to move vertically on "tracks," depicted by the shaded areas in Figure 3. At the beginning of the task the two objects were centered on the targets, but tended to drift away from the target lines toward either the upper or lower edges of the window in response to random perturbations. The location of each object was determined by a fixed X coordinate (horizontal position) and a variable Y coordinate (vertical position). With each iteration of the main program loop, a random perturbation of plus or minus one pixel was added to the Y-coordinate. The result was a rapid, irregular "jiggle" in object position. Once an object was displaced from the starting position on the target line, it moved more and more rapidly away from the target. The mechanical analog was an object rolling downhill, with the slope of the hill getting steeper as the horizontal distance from the top increased. The coefficients governing the "steepness" of each virtual "hill" were selected by the experimenter, and were the same for all subjects. Figure 3 shows the circle as having moved off the target and come to rest at the bottom of the window; the square has come to rest at the top. Either object could have come to rest in either position, absent participant input intended to keep them on target.

Figure 3 About Here

Figures 4 and 5 explain the origin of the "controller" terminology. Moving the controller vertically caused the circle to move in the same direction, with an apparent "force" that was proportional to the difference in the Y-coordinates of the two objects. That is, the further it was displaced away from the circle, the more strongly the controller "pulled" it. The mechanical analogy was a rubber band connecting the circle and the controller, but with an important difference; only the vertical component of the distance, denoted Y1 in Figure 4, had any effect. The horizontal distance between the circle and the controller had no effect. The square was controlled in a similar fashion by displacing the controller horizontally. Moving the controller to the left of center impelled the square toward the top of the screen with a "force" proportional to the sum of distances X1 and X2. Moving the controller to the right had the opposite effect. Just as the horizontal position of the controller had no effect on the circle, the vertical position of the controller had no effect on the square. The coefficients determining how strongly the controller "attracted" the square and the circle were selected by the experimenter, and were the same for all participants.

Figures 4 and 5 About Here

An object was scored "on target" whenever the target line intersected some portion of the object's area. The participant manipulated the mouse as necessary to maintain both objects on target simultaneously, as shown in Figure 6. While this condition prevailed the participant was making progress toward successful completion of the task.

Figure 6 About Here

For each participant, the program maintained a record of two elapsed times; a total time and a time on target. At the beginning of the task both the "total time (TT) clock" and the "time on target (TOT) clock" were set to zero. The total elapsed time of each clock increased under specified conditions, and was not reset to zero until the task was complete. The TT clock ran continuously whenever the task was executing; i.e., when the graphical objects were in motion and subject to participant control. When both objects were on target (Figure 6), the TOT clock ran; if either object moved off target, the TOT clock stopped until the on-target condition was re-established. The task was complete when either the TOT clock reached one minute or the TT clock reached 30 minutes. If the first condition was obtained before the second, the exercise was graded as a success; otherwise, it was a failure.

Two scores were computed, an absolute score and a quality score. The absolute score was the number of seconds on the TOT clock, and was a direct indicator of the subject's progress. It ranged from an initial value of 0 to a final value of 60. The quality score was the ratio of TOT to TT, and represented participant skill. The quality score ranged from a minimum of zero (not achievable in practice, since the task began with both objects on target) and a maximum of one (theoretically achievable, but unlikely).

As noted above, the participant was able to interrupt execution of the task by clicking the mouse. While the task was in the interrupt mode (i.e., display frozen, clocks not running) the participant was able to navigate a hierarchy of menus and obtain feedback. Menu items were presented as point-and-click buttons (Figure 7; the participant indicated his or her choice by placing the cursor arrow in the box to the left of the item, and clicking the mouse.

Figure 7 About Here

Figure 8 depicts task execution in flowchart form[3]. The heavy rectangles with rounded corners are screens. The experimenter entered the participant's identification (Social Security account number) and verified which one of the four experimental manipulation (described below) was in effect. The participant initiated task execution by reading a message explaining the use of the point-and-click protocol, and indicating his or her understanding by selecting a button labeled "Continue" (Screen SC). The participant viewed a series of introductory screens (I1 through I8) explaining how the task was to be performed and scored. The participant had the option of repeating individual pages, or the entire introduction, as often as desired; the number of iterations he or she requested was recorded. When satisfied with his or her understanding of the introductory material the participant clicked on the "Continue to task" button. After clicking in the start-task (ST) screen, the participant attempted to control the objects on the task (T) screen. He or she interrupted task execution as desired by clicking the mouse.

Figure 8 About Here

Upon interrupting task execution, the participant was presented with the main feedback menu (Screen MM: Figure 8, second page). At this point the participant had three options; returning to the task, requesting outcome feedback in the form of performance information, or requesting "process feedback" in the form of lessons about how to improve performance. If outcome feedback was selected (screen OUT), the participant was presented with the total TOT accrued, the remaining TOT required for task completion, and his or her quality score at that point. When finished with the OUT screen the participant returned to the main feedback menu (MM). From the main menu the P was able to return to the task, or to request additional feedback.

If the participant requested process feedback, a more elaborate protocol was utilized. Ideally, process feedback is information about how one <u>should</u> perform a task,

[3] A complete program listing, including details of the screen displays, is available from the author.

based upon one's actual performance. The present program was not sophisticated enough to diagnose individual performance, and therefore could not offer true process feedback. Instead, the program presented a form of simulated process feedback. Although it purported to address each participant's individual performance, the text was in fact identical for all participants. This subterfuge - i.e., that the feedback presented was in fact keyed to the participant 's performance - seemed reasonable because of the task's novelty. It was decided that participants would probably find it difficult to generate, on their own, any process feedback detailed enough to contradict the information presented.

All subjects received the same items of process feedback in the same sequence, thus eliminating the relative usefulness of the items as a source of variance. The first time process feedback was selected, the participant is presented with lesson 1 (L1); the second time process feedback was selected the participant was presented with lesson 2 (L2), with the option to review L1, etc. If the participant revisited the process feedback option more than four times, he or she was immediately given a menu of all four lessons, with instructions to select the one desired. This mode of presentation fostered the illusion that the program was "prescribing" feedback in response to individual performance. Such an illusion would have been hard to sustain if participants had been given free access to all feedback items at any point during task execution.

The process feedback procedure is shown on the second page of Figure 8, beginning with the main menu (MM) screen. The participant requested process feedback from the menu. The program determined which request this was (the first, second, etc.) and presented the appropriate lesson. Lesson 1 (L1) addressed the correlation between cursor and controller motions, and offered a practical tip on how to hold the mouse. L2 discussed the correlation between controller and circle motion, and L3 covered the correlation between the controller and the square. L4 emphasized the importance of division of attention, and gave a practical tip on how to effectively move one's gaze around the display. After reading the narrative of each lesson, the participant viewed a demonstration or performed an exercise; each was begun by clicking the mouse in a start screen (D1S through D4S) and terminated by clicking in the demo screen (D1 through D4). After each demonstration the program presented an appropriate menu (M1 through M5). Each menu offered the option of returning to the main menu or reviewing previous items of process feedback.

Each request for feedback was recorded by the program, along with the TT and TOT associated with each instance. After the participant had viewed feedback items, he or

she returned to the task via the resume-task (RT) screen. When TT reached 30 minutes or TOT reached 60 seconds the task terminated.

Task manipulations

Using the survey results, participants were rank-ordered on external propensity, internal propensity, and internal ability. Those above the median in the rank order were considered "high" in that attribute; those below the median were considered "low." The high and low levels of each of the three attributes were crossed to define eight groups of participants. To the greatest extent possible, participants in each of the eight participant groups were randomly assigned in equal numbers to the manipulations using the procedure described above. It should be noted that these individual differences were not dichotomized for purposes of analysis; the assignment procedure was intended to distribute individual differences as evenly as possible among the experimental manipulations. The assignment procedure was unavoidably flawed, since it could not be applied to all the participants as a single group. Participants became available in three separate cohorts, over the period of two academic quarters; each cohort was separately tested and divided into the eight treatment groups, so that they could be scheduled for the laboratory task.

Social presence and the supposed nature of the evaluation were manipulated using a 2x2 between-subjects design involving two levels of observer presence and two different task descriptions. The observer conditions were (1) private (e.g., the participant was alone in the room) and (2) observed (the task was conducted with an experimenter, microphone, and TV camera present.). The task descriptions were (1) an evaluation of the computer program, as an example of "computer-assisted instruction", and (2) an evaluation of the participant, via a "computerized aptitude test." As discussed above, feedback eliciting was expected to covary with these manipulations, with the greatest differences occurring between the 1-1 cell of the design (manipulation I) and the 2-2 cell (manipulation IV, below). It was expected that the 1-2 and 2-1 treatments (manipulations II and III) would produce intermediate effects, although no hypotheses concerning rank order were offered.

A cubicle in the Organizational Behavior Laboratory of the Georgia Tech School of Management was used (Figure 9). The participant initially sat at a small table in the corner and completed the prequiz. After completing the prequiz and depositing it in the box provided, he / she turned the chair to face a 12 inch color monitor and a computer mouse, both connected to a Macintosh PowerBook 160 computer. The PowerBook was set up back-to-back with the monitor, facing the experimenter's station. The participant was able

to perform the experimental task, to include navigating the feedback menus, by using the mouse alone. All procedures requiring keyboard entries were performed by the experimenter, who was seated in front of the computer.

Figure 9 About Here

Different room arrangements were used to change the experimental treatment of observer presence are shown in Figure 9. In the "Observed" condition, the experimenter was present throughout. The participant was recorded on videotape while working the task, and was able to see his / her image in quarter profile on the video monitor. A microphone connected to the VCR was suspended over the participant's station. In addition, the two-way mirror between the cubicle and the room next door was uncovered, suggesting the possibility that surreptitious observers may have been present. In the "Private" condition, the video camera was turned off, and covered with a black plastic trash sack. The television monitor displayed a blank screen. The microphone was taken down and covered along with the camera, and a large sheet of paper was hung over the 2-way mirror using tabs of hook-and-pile tape.

The treatment involving the purpose of the task was administered by induced by showing the participant one of four short video presentations. The presentations were professionally prerecorded using a faculty member as the "actor," and explained the putative nature of the task (i.e., evaluating either the program or the participant). In addition, the messages provided the rationale for the presence or absence of the experimenter, and for videotaping the task. The main features of each manipulation, as explained in each script, are shown in Table 1.

Table 1 About Here

The Postquiz

Following completion of the task, but before being debriefed on his or her performance, each participant completed the postquiz. This was accomplished at the table in the corner with the privacy partition, where he or she had filled out the prequiz earlier (Figure 9); following completion, the postquiz was deposited in the same box as the prequiz. Physically, the postquiz consisted of a nonreusable paper form with multiple-

choice responses that the participant marked with either pen or pencil. Data were transcribed from the form to a data file by the experimenter.

Items on the postquiz assessed the participant's state anxiety while working on the task, plus his or her impressions of feedback availability and usefulness, and a self-evaluation of task performance.

Measures

Variables quantifying the two criteria, total feedback eliciting and the ratio of outcome to total feedback eliciting, were calculated from data recorded by the computer program. Two operationalizations were formulated for each criterion. The first measure was based on the number of times, or instances, that the participant had requested feedback; the second, on the number of seconds he or she had spent perusing the feedback items after requesting them. Each measure was designed to compensate for possible weaknesses in the other. For example, two people may ask for feedback the same number of times. The first may be much more concerned with the information presented, and spend more time studying it; he or she would therefore be "eliciting" more feedback than the second, in that he or she would be attempting to extract more performance-relevant information from the data presented. The difference would appear in time data, but not in a count of feedback-eliciting instances. Conversely, two subjects may spend the same amount of time attending to feedback, but the second may elicit it more often, thus engaging in more proactive feedback seeking. This difference in behavior would appear as a difference in the number of feedback-seeking instances, but not as a difference in time.

The Total Feedback-eliciting Count (TFCNT) was the total number of times the participant requested either outcome feedback or process feedback. The Outcome Eliciting Ratio based on Count (OERCNT) was the number of times the participant requested outcome feedback, divided by his or her TFCNT. This measure was only computed for cases having TFCNT not equal to zero. The Total Feedback-eliciting based on Seconds (TFSEC) was the total number of seconds the participant spent studying either outcome or process feedback items. The Outcome Eliciting Ratio based on Seconds (OERSEC) was the number of seconds spent studying outcome feedback, divided by TFSEC (if not equal to zero).

Performance (PERF) was operationalized as the number of seconds the participant kept the screen objects on target (TOT), divided by the total number of seconds (TT) he or she was engaged in the task. As noted in the task description above, 60 seconds

TOT was required for successful completion; maximum task time was 30 minutes, or 1800 seconds. Maximum possible PERF was 1.00, corresponding to 60 seconds on target and zero seconds off-target. Minimum possible PERF was 0.0001, which was only observed during task development when the experimenter deliberately moved the objects off-target as quickly as possible, and kept them there for the full 30 minutes. None of the participants came close to either of these two extremes.

External Propensity (EXT) was measured during the preliminary survey, utilizing a six-item scale developed by Herold and Parsons (1985). A representative item was, "It is very important for me to know what people think of my work." Participants responded to items utilizing a five-point scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Task-specific Internal Ability (TIA) was measured during the preliminary survey. Items originally developed by Herold and Parsons (1985) were modified to reflect the respondent's confidence in his or her ability to self-generate feedback with respect to computer-based tasks. A representative item was, "When doing a task on a computer, I can usually tell early on if it is going to turn out OK." Participants responded to items utilizing a five-point scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Task Familiarity (TFAM) was measured during the pre-experiment quiz (prequiz). Measuring TFAM was problematical, since the task was designed to be as novel as possible and participant familiarity was, ideally, very low. In addition, it was not possible to ask, "How often have you done something like this before?" on the prequiz without first describing the task in some detail. Offering such a description was deemed to be undesirable, since it may have affected subsequent feedback-eliciting, or created preconceptions that may have affected performance. However, one antecedent factor related to task familiarity was identified; namely, participants' experience with the graphical user interface (GUI), including the use of a mouse to control computer tasks. Having such experience was expected to increase participants' comfort levels as well as performance, and therefore affect feedback-eliciting. The participants were asked, "How often do you use a computer with a graphical 'point-and-click' operating system, like the Mac, or a DOS-based machine running Windows?" Participants responded to the item utilizing a six-point scale ranging from 6 ("Daily") to 2 ("Almost never") and 1 ("I don't know what you're talking about.")

Internal Propensity (INT) was measured during the preliminary survey, utilizing a six-item scale developed by Herold and Parsons (1985). A representative item was, "How other people view my work is not as important as how I view my own work." Participants

responded to items utilizing a five-point scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Self-Esteem (SE) was also measured during the preliminary survey, utilizing a ten-item scale adapted from Rosenberg (1965). Two items were "I feel I have a number of good qualities" and "I certainly feel useless at times" (reverse scored). Participants responded to items utilizing a five-point scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Internal-External Locus of Control (LOC) was measured during the preliminary survey, using fifteen items developed by Rotter (1966). A representative item was, "Most people don't realize the extent to which their lives are controlled by accidental happenings." Participants responded to items using a five-point scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Tolerance for Ambiguity (TA) was measured during the preliminary survey using twelve items adapted from Norton (1975) and Ashford and Cummings (1985). Two representative items were: "Once I start a task, I don't like to start another task until I finish the first one." "Before any important job, I must know how long it will take." Participants responded to items utilizing a five-point scale ranging from 1 (strongly disagree) to 5 (strongly agree), but all the items were later reverse-scored to permit a more straightforward interpretation; i.e. tolerance for, rather than intolerance for, ambiguity.

Need for Achievement (NAch) (McClelland, 1961), was measured using ten items adapted from Steers and Brownstein (1976). One representative item was, "I do my best work when my job assignments are fairly difficult." Participants responded to items utilizing a five-point scale ranging from 1 (strongly disagree) to 5 (strongly agree). Ever since McClelland (1961) proposed the construct, investigators have debated the relative merits of the Thematic Apperception Test (TAT) versus questionnaires as instruments for assessing nAch. A recent meat-analysis by Spangle (1992) has shown that questionnaire measures of nAch are significantly correlated with various types of outcomes, ranging from attitude measures to income and occupational success. Although TAT scores typically produce slightly higher correlations with outcomes than questionnaire scores (Spangler, 1992), the logistical difficulties involved in administering and scoring the TAT for a large number of participants dictated the use of a questionnaire-type instrument in the present study.

The two manipulated variables were social presence (SOCIAL) and the evaluation target (TARGET). SOCIAL was coded 0 for participants who performed the task in

private, and 1 for those who had an observer, VCR, open window, etc. as explained above. TARGET was coded 0 for participants who were told they were helping gather data to evaluate the task as "an example of computer assisted instruction" -- that is, for participants who were not the targets of evaluation. TARGET was coded 1 for those who were told that they were performing "tests of psychomotor, intellectual, and analytical ability" that would "predict your academic performance..."

Levels of SOCIAL and TARGET were hypothesized to affect feedback seeking through the mechanism of arousal. Therefore, it was deemed essential to incorporate a measure of arousal into the experiment as a manipulation check. Using a physiological measure of arousal such as pupil dilation, pulse rate, or galvanic skin response was not feasible, which forced reliance upon a questionnaire-type self report. A self-report of anxiety was adopted. It was recognized that arousal and anxiety are not isomorphic; sexual contact, for example, may produce arousal, but not (necessarily) anxiety. On the other hand, it was reasonable to assume that increased arousal due to performing a stressful task, being observed, and / or being evaluated, would be reported as elevated levels of arousal. The instrument chosen was the State Anxiety Inventory of Spielberger et al (1977). The 20 items were rewritten from the present to the past tense. Participants were asked to reflect on the way they felt while they were working on the task, which they had just completed. Instead of the original four-point scale, participants responded to items utilizing a five-point scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Internal consistencies for multi-item scales are shown in Table 2. As noted above, the survey and the laboratory experiment were presented to participants as separate experiments, and participation in one did not require participation in the other. This resulted in different sample sizes (N) for survey participants, laboratory participants, and for those taking part in both. The N varied slightly among the scales within each group due to the effects of missing data; if one item was omitted, inadvertently or otherwise, the scale to which it belonged could not be calculated. Not unexpectedly, the most restricted sample (survey plus lab participants) produced the smallest internal consistencies; external propensity, internal propensity, tolerance for ambiguity, and need for achievement all had values of Cronbach's alpha less than 0.70. A discussion of the effects these weak internal consistencies may have had on tests of hypotheses will be presented later.

Table 2 About Here

ANALYSES

Because the most common statistical tests assume normal distribution, all of the variables were tested for that property using the Kolmogorov-Smirnoff goodness of fit test. The *p* statistic represents the probability that the data for that scale were drawn from a normally distributed population. Two operationalizations of the criteria (total feedback-eliciting count and the outcome eliciting ratio in seconds) and three predictors (performance, external propensity and task familiarity) were non-normal at the 0.05 level of significance. This suggested the appropriateness of including nonparametric tests in the analyses. Therefore, both Pearson and Spearman correlation coefficients were calculated.

Regression analysis was used for testing both main effects and interaction hypotheses. The 12 predictors were regressed as a block on the two feedback-eliciting variables (TFCNT and TFSEC), using ordinary least-squares regression. Listwise deletion reduced the number of cases from 129 to 117; any case having a missing value for any of the criteria or predictors was omitted from the analysis. The same procedure was utilized for the two outcome-ratio variables (OERCNT and OERSEC), and their predictors. In the cases of OERCNT and OERSEC, listwise deletion plus the requirements that TFCNT and TFSEC be greater than zero, reduced the number of cases to 109.

Interactions were examined using the following procedure. The 12 predictors of TFCNT and TFSEC were crosstabulated, paired and multiplied to produce a list of 66 interaction terms. Due to power constraints these could not all be tested simultaneously. Regressing 12 predictors and 66 interaction terms with only 117 subjects would have produced an analysis with approximately 1.4 subjects per regression term, which would have been meaningless. Instead, the interaction terms were entered following the block of predictors using a stepwise procedure with an entry criterion of p=.05 and an elimination criterion of p = .10 (Cohen & Cohen, 1983). This procedure yielded a list of candidate regression terms with statistically significant beta weights.

The 0.05 criterion, applied at each step of the stepwise regression, obviously did not limit the overall probability of Type 1 error to 0.05. Consider a hypothetical list of 66 predictors, none of which is significantly related to some criterion in the population. If the probability of Type 1 error with respect to any one of the predictors was 0.05, then the binomial rule shows that the probability of rejecting all of them (the correct outcome) was only 0.033. For this reason, the candidate regression terms were subjected to the

following additional test. The criterion was regressed on the two predictors making up each interaction, as a first hierarchical step. The interaction term was then entered as a second step. If the change in F following the entry of the interaction was significant, then the interaction was retained for interpretation. Finally, the criteria were regressed on the block of original predictors, followed by a block of all the retained interaction terms. The same procedure for analyzing interactions was applied to OERCNT, OERSEC, and their predictors.

Significant interactions were interpreted graphically using a procedure suggested by Cohen and Cohen (1983). An equation was calculated by regressing each criterion on the interaction term and its two constituents. Four values of the criterion were then calculated, using the coefficients of the equation and two values for each of the two predictors; these values consisted of the mean of each predictor, plus and minus one standard deviation. Two regression lines were then plotted, one for the high and one for the low value of the first predictor, with end points determined by the high and the low value of the second predictor.

RESULTS

Examination of the Pearson and Spearman rank-order correlation coefficients between criteria and predictors (in the boxes, Table 3) suggested that the latter may permit more sensitive tests of the experimental hypotheses, since the absolute values of the Spearman coefficients appeared to be larger than their Pearson equivalents. This apparent difference was confirmed by applying Fisher's r-to-z transformation to both sets and comparing them using a paired t-test. The means (Pearson mean = -0.03, Spearman mean = -0.06) were significantly different ($t = 2.19$, $df = 47$, $p = 0.035$).

Table 3 About Here

It also appeared that the size of the difference between the Pearson and Spearman correlation coefficients were related to the degree by which the correlated variables departed from normality. For example, both TFCNT and PERF are significantly non-normal by the Kolmogorov-Smirnoff (K-S) test. For this pair, the absolute difference between the Pearson and Spearman coefficients was 0.14. On the other hand, both TFSEC and STANX were not significantly non-normal, and the difference between their coefficients

was only 0.03. Attempts to demonstrate this relationship statistically by correlating coefficient differences (as z-scores) with mean K-S z-scores (indicating divergence from normality) failed, however, due to the preponderance of small, mostly non-significant correlations in both sets.

More to the point, three hypothesized relations had Spearman coefficients significant at the 0.05 level and in the hypothesized directions, while their Pearson counterparts were not significant (Table 3). In addition, there were two significant Spearman correlations for which hypotheses were not offered (between SOCIAL and both OERCNT and OERSEC.) The converse was not true; all significant Pearson coefficients had significant Spearman counterparts. Based on these considerations it was decided to use the Spearman coefficients for hypothesis testing, as better representing the relationships among non-normal variables.

Since only a small subset of the correlations was significant, the first question posed of the data concerned the global null hypothesis; i.e., whether the appearance of this many significant correlations could reasonably be attributed to chance. Thirty-four of the 48 correlations represented experimental hypotheses concerning feedback eliciting, while 14 (e.g., EXT with OERCNT) did not. Of the 34 correlations that "mattered," nine were significant at the 0.05 level. If one assumes random data and sets the probability of Type 1 error at 0.05, then the probability[4] of nine spuriously significant correlations out of 34 is 2.84×10^{-5}.

In addition, 21 of the 34 correlations were in the hypothesized direction; i.e., had the "correct" algebraic sign. Using the previous argument and assuming random data, the probability of obtaining a "correct" sign by chance is 0.50, and the probability of obtaining 21 spuriously correct signs out of 34 is 0.0540. Clearly the Spearman correlations had a story to tell.

The significant negative correlations between PERF and both TFCNT and TFSEC provided support for H1. Higher levels of performance are correlated with lower levels of feedback-eliciting for both operationalizations of the behavior; both the number of instances of eliciting, and the total time spent eliciting. Similarly, the negative correlations between PERF and both OERCNT and OERSEC provided support for H2. Higher levels of performance were correlated with lower levels of OER, where OER was the ratio of

[4] Binomial probability P(r) with number of "hits" r=9, number of trials n =34, and the probability of a "hit" per trial p = 0.05

outcome feedback-eliciting to total feedback-eliciting. Again, the hypothesized relationship was obtained for both operationalizations of OER, both OER calculated in terms of instances (OERCNT), and OER calculated in terms of time (OERSEC).

The Spearman correlations provided no support for H3. In the last column of the top half of Table 3, a correlation of 0.32 ($p<=.01$) provided support for the hypothesized relationship between task familiarity and task-specific internal ability (H4a). Those reporting more experience with the graphical user interface had more confidence in their ability to self-generate valid feedback when engaged in computer-related tasks. The significant negative correlation between TIA and TFSEC provided partial support for H4b; higher levels of task-specific internal ability were associated with lower levels of feedback-eliciting, but only when the behavior was operationalized in terms of time. Similarly, the negative correlation of TFAM with TFSEC provided partial support for H4c. Higher levels of task familiarity were associated with lower levels of eliciting, but only in terms of the time spent eliciting, not the number of discrete eliciting behaviors.

The significant negative correlation between INT and TFCNT lent partial support to H5a; higher levels of internal propensity were related to lower levels of eliciting, but only to the number of eliciting instances, not the to time spent. H5b was not tested by correlations, but rather by examining the interaction between TIA and INT (below). H6a through 10d received no support from the Spearman correlation coefficients.

It was hypothesized that state anxiety (STANX), as a measure of arousal, would be correlated with the experimental manipulations of social presence (SOCIAL) and evaluation target (TARGET). Neither correlation was significant. Further, neither SOCIAL nor TARGET was correlated with either TFCNT or TFSEC, thus failing to support H10a and H10b. This was perplexing, since STANX did vary among the participants, and was found to be correlated with both TFCNT and TFSEC, although not in the direction hypothesized. This outcome will be discussed at greater length below.

Regression Results

Multivariate regression was used to examine the relationships between criteria and predictors while partialling out the effects of shared variance. The results of regressing the predictors on the measures representing total feedback-eliciting, and on those representing the ratio of outcome eliciting to total eliciting, will be presented separately.

The results of regressing the predictors on the two operationalizations of total feedback-eliciting, TFCNT and TFSEC, are shown in Table 4. The significant Spearman

correlations between STANX and the criteria suggested that STANX be included as a predictor. H1 was supported by the significant negative regression weights associated with the performance variable, PERF. H4c was only partially supported, in that task familiarity (TFAM) was a significant negative predictor to the time spent eliciting feedback (TFSEC), but not the number of feedback-seeking instances (TFCNT). The data indicate that participants having more experience with the graphical point-and-click protocol spent less time perusing feedback items, but did not necessarily request fewer of them. As before, state anxiety (STANX) was significantly related to both TFCNT and TFSEC in directions contradicting H10e. Surprisingly, need for achievement (NAch) emerged as a significant predictor of both TFCNT and TFSEC, a relationship not suggested by the Spearman correlations, and even more unexpected considering the measure's low coefficient alpha (0.606, Table 2).

Table 4 About Here

Six significant interactions were identified between predictors of TFCNT and TFSEC. For the criterion TFCNT, three interactions contributed an additional 16 percent to explained variance (equation 2, bottom of Table 4). For TFSEC, three interactions contributed an additional 18 percent to explained variance (equation 4, bottom of Table 4). The only hypothesis framed as an interaction between two predictors, TIA and INT (H5b), was not supported. The first interactions to enter utilizing the stepwise procedure involved two of the main effects, performance and state anxiety (Figure 10). For both TFCNT and TFSEC, lower levels of performance predicted higher levels of eliciting, but the effect was more pronounced for those reporting higher levels of anxiety.

Figure 10 About Here

The interactions between performance and tolerance for ambiguity as predictors of total feedback eliciting are shown in Figure 11. Participants engaged in more feedback eliciting at lower levels of performance, but the effect was more pronounced for those who reported high levels of ambiguity tolerance. As Figure 11 shows, the phenomenon appears for both operationalizations of total feedback eliciting, TFCNT and TFSEC.

Figure 11 About Here

Figures 12 and 13 display interactions between internal propensity and task familiarity, and between internal propensity and social presence. More restricted interpretations are required than before, because each of the interactions appears with only one of the two variables representing total feedback-eliciting. The INT x TFAM interaction (Figure 12) has TFCNT as the criterion; the INT x SOCIAL interaction has TFSEC as the criterion.

Figures 12 and 13 About Here

Figure 12 shows that participants who described themselves as having a high degree of familiarity with tasks of this type (i.e., utilizing the graphical user interface) demonstrated the hypothesized relationship between internal propensity and feedback eliciting; those with low INT elicited more, those with high INT elicited less. Those with low familiarity, on the other hand, demonstrated the opposite relationship: those with low INT elicited less, and those with high INT elicited more. This result suggests that the effects of internal propensity are moderated by one's need for information.

Figure 13 suggests that the effect of internal propensity on feedback eliciting is strongly conditioned by the presence of another person, even when that person is not the source of feedback. When an observer was present, the effect of INT on eliciting conformed with hypothesis; lower levels of INT resulted in more eliciting, higher levels of INT with less. When participants performed the task in private, however, the relationship between INT and eliciting was much weaker; in addition, the sense of the relationship was reversed. The result suggests that the behavioral consequences of INT, which is an expressed preference for self-generated feedback, are manifested when a witness to external feedback-seeking is present, but otherwise are not.

The results of regressing predictors on the ratio of outcome to total feedback eliciting (OER) is shown in Table 5. As before, the construct was represented by two variables, one based on the number of times feedback was requested (OERCNT), the other based on the number of seconds spent studying feedback items (OERSEC). Only H2

received partial support; performance (PERF) was a significant predictor of OERCNT, but not OERSEC. There were no significant interactions.

Table 5 About Here

DISCUSSION

This research had two objectives. The first was to investigate selected antecedents of feedback eliciting, using objective measures of eliciting instead of self-reports. The second was to see if certain factors affected the relative frequencies of outcome and process feedback eliciting. We will address the two objectives separately.

The first objective was motivated by methodological concerns. Many previous studies had reported correlations between questionnaire-based individual differences, such as tolerance for ambiguity, and self-reports of feedback-seeking. Even when techniques for eliminating common-method variance were utilized, the fact remained that both data came from the same source -- the individual. In effect, people were being asked who they were and what they did. Some degree of spurious correlation between the two responses seems predictable. For example, it seems inevitable that the student pilots who reported a high degree of external propensity also would have reported higher levels of feedback-seeking behavior (Fedor et al, 1992), even if their behavior was in fact the same as everyone else's. If one has a need for external feedback, and is therefore highly aware of one's own feedback-seeking behaviors, then it is reasonable to see how one could retrospectively overestimate the frequency of such behaviors. One can expect that self-reported individual differences would have lower correlations with objective measures of "real" behavior than with self-reports of behavior.

A cursory examination of Table 6 below tends support such an expectation. Let us concentrate on the results of multivariate regression, since it provides a stronger test of the hypotheses than correlation. Only one trait-type individual difference, need for achievement, had significant regression weights with respect to both measures of feedback eliciting, and only after the effects of performance, task familiarity, and state anxiety had been partialled out of every other predictor. The feature common to performance, task

familiarity and state anxiety is that they all were related to the task, or to the individual's relationship with the task. They were not pure "individual differences." Just the same, one cannot discount the effect of individual differences in favor of situational factors. First, NAch was an individual difference, and was also an unambiguous predictor of feedback eliciting. Second, Table 6 is a summary, and suppresses a lot of the inconclusive but suggestive details reported earlier. Revisiting Table 3, one counts 14 Spearman correlations between "pure" individual differences (EXT, TIA, INT, SE, LOC, TA and NAch) and the two operationalizations of total feedback eliciting (TFCNT and TFSEC). Of these, 11 are in the hypothesized directions. Given a 0.50 probability of obtaining a "correct" sign by chance, the likelihood of such an outcome is 0.02. But, on the other hand, only two of the 14 correlations were statistically significant. Part of the correlational weakness may have been due to the low internal consistencies noted above (Table 2), but this cannot be the whole story; NAch, the only individual difference to achieve significant beta weights under multivariate regression, had the lowest coefficient alpha of the group (0.606). In anthropomorphic terms, one has the impression of individual differences struggling to express themselves as behavior, but being overwhelmed by the exigencies of the task. This impression is supported by Mischel (1968), who asserted that in unfamiliar or stressful situations, individual differences are subordinate to situational factors as determinants of behavior.

Table 6 About Here

Within the context of this particular task, performance and state anxiety (and to a lesser extent task familiarity) dominated feedback eliciting. Is it possible to say anything about the antecedents of these factors? Fortunately, some additional data were collected that permitted exploratory analyses. Before discussing these analyses, though, we will examine the original regression results more closely.

Although H1 was the most strongly supported, a caveat must be registered. Performance was operationalized as time-on-target, in seconds, divided by the total time the participant worked on the task. This total task time was the number of seconds the screen objects were in motion, and excluded the time the participant spent selecting items from menus, studying outcome feedback, reading lessons, and viewing demonstrations.

Feedback eliciting was operationalized in terms of (1) the number of times the participant requested feedback, and (2) the number of seconds he or she spend perusing the feedback. All of these measures are related in the following manner. Total task time, the time spent perusing feedback, and the total time available to request feedback, were all subsets of the total time the participant spent in front of the computer. Two things argue against the possibility that the observed regression weights are spurious. First, neither one is suspiciously large; second, one would expect that the operationalization of eliciting in terms of time (TFSEC) would be more susceptible to inflation than the operationalization in terms of instances (TFCNT). In fact, the regression weight of TFSEC is smaller than that of TFCNT. It seems safe to conclude that the regression weights support the hypothesis, as manifestations of the ways that participants allocated their time among the various task activities.

Hypothesis 4c was only weakly supported by regression analysis. Task familiarity was operationalized as the participant's self-reported familiarity with the graphical user interface, and this may have been a weak measure. As noted above, obtaining a highly veridical measure of task familiarity would have required giving participants a detailed preview of the task, which would have had an effect on feedback-seeking. Task familiarity had a small and marginally significant beta when regressed on TFSEC, the time measure, but not on TFCNT, the count measure. It seems possible that participants having greater familiarity with the point-and-click protocol (high TFAM) may have spent less time transitioning between lessons and demonstrations. This time that was included in the number of seconds spent performing process feedback-seeking. Thus, simple skill using the computer mouse may have produced the observed relationship between TFAM and TFSEC, and not the level of feedback eliciting. This result for H4c should be viewed with caution.

As previously noted, the support for H9a came as a surprise. Need for achievement had weak internal consistency, and no significant Spearman correlations with either operationalization of total feedback eliciting. When the effects of performance and apprehension were partialled out, however, need for achievement contributed a significant increment of explained variance. This can be understood in light of Weiner's (1978) theory; people with high need for achievement actively seek feedback, either through selecting tasks with medium difficulty (and high diagnosticity), or through a higher level of active feedback elicitation, as this result indicates.

One of the most surprising findings was the failure of experimentally manipulated levels of social presence and evaluation to predict differences in feedback eliciting. Neither social presence (SOCIAL) nor evaluation target (TARGET) was correlated with either measure of total feedback eliciting (TFCNT and TFSEC; Table 3), nor did they have significant regression weights (Table 4). Only social presence appeared as a interaction component, interacting with internal propensity as a predictor of total feedback eliciting in seconds (see above). State anxiety, which did have significant regression weights with TFCNT and TFSEC (Table 5), was not correlated with either SOCIAL or TARGET as hypothesized. Additional analyses were undertaken in an attempt to understand these results.

First, feedback eliciting was reexamined under the conditions most likely to have produced differences. It was assumed that the strongest effects would appear as differences in behavior between the 1-1 cell (manipulation I) and the 2-2 cell (manipulation IV) of the design: to recapitulate, the former was the condition in which participants were supposedly gathering data to evaluate a computer task, without an observer present; in the latter, they were supposedly performing a computerized aptitude test with an observer present. A new variable called STRESS was defined, coded 0 in the first manipulation (n = 41) and 1 in the second (n = 29); participants in the manipulation II and III conditions cells were disregarded. Analysis of variance revealed no significant relationship between STRESS and state anxiety (F=.528; df = 1, 68; p= .470), nor between STRESS and any of the four feedback eliciting criteria. For the latter, univariate F-tests with df = 1, 62 produced the following statistics: for TFCNT, F=.000 (p=.977); for TFSEC, F=.411 (p=.524); for OERCNT, F=1.293 (p=.260); and for OERSEC, F=1.621 (p=.208). In sum, eliminating ambiguous manipulation conditions and concentrating on those believed to be strongest revealed no effects with respect to the criteria.

Another analysis examined the possibility that the zero-order correlations between the predictors and the criteria were significantly different between the two STRESS conditions. For example, need for achievement may have been more highly correlated with feedback eliciting in the STRESS = 2 condition, since behavior may have been more influenced by the participants' innate achievement needs when they were confronted by an observer and told that they were being evaluated. In other words, it was thought that the experimental conditions may have differentially "turned on" various individual differences as predictors of behavior, resulting in statistically significant differences between their correlations in the different conditions.

Pearson correlations between the predictors and the criteria were calculated for both STRESS conditions. The correlations were converted to z-scores using Fisher's r-to-z transformation[5]. Next, the differences between the z-scores, divided by the standard error of the difference, were computed (McNemar, 1969). The value of p associated with each z-score difference indicated the probability of finding a difference of that magnitude when correlations in the population were actually the same under both conditions. Only four significant differences were found. They involved the correlations between task familiarity (TFAM) and the two measures of the outcome eliciting ratio (OERCNT and OERSEC), and the correlations between internal propensity (INT) and the two measures of total feedback eliciting (TFCNT and TFSEC).

The correlation between task familiarity and OERCNT in the STRESS = 0 condition was -0.2662 (p = .117); in the STRESS = 1 condition, it was +0.1514 (p = .442). The z-score difference between the correlations was 1.60 (p = .0548). The correlation between task familiarity and OERSEC in the STRESS = 0 condition was -0.2274 (p = .182); in the STRESS = 1 condition, it was +0.2157 (p = .270). The z-score difference between the correlations was 1.70 (p = .0446). Thus, a treatment effect appeared for the correlations between task familiarity and both measures of the outcome eliciting ratio. Under conditions of lower STRESS, in manipulation I, there were negative correlations between task familiarity and the outcome eliciting ratio. Those who reported high task familiarity elicited less outcome feedback as a proportion of total feedback. Under conditions of higher stress, in manipulation IV, the relationship was reversed. This result is difficult to interpret, but it argues for the existence of an effect due to either social presence, or evaluation, or both jointly, upon the relationship between task familiarity and the type of feedback sought.

The STRESS treatment condition created a difference between correlations similar to the one above. The correlation between internal propensity and TFCNT in the STRESS = 0 condition was -0.080 (p = .619); in the STRESS = 1 condition, it was -0.560 (p = .001) The z-score difference between the correlations was -2.19 (p = .0143). The correlation between internal propensity and TFSEC in the STRESS = 0 condition was -0.041 (p = .800); in the STRESS = 1 condition, it was -0.625 (p = .000). The z-score difference between the correlations was -2.75 (p = .0446). Again, a treatment effect

---

5 Pearson correlations were used instead of Spearman rank-order correlations, because of uncertainty concerning the applicability of the r-to-z transformation to the latter.

appeared for both measures of the criterion: in this case, total feedback seeking. Under conditions of lower STRESS, in manipulation I, there were negative correlations between internal propensity and total feedback eliciting, but they were small and not statistically significant. Under conditions of higher stress, in manipulation IV, the relationship was stronger. This result suggests that internal propensity has a social component; that is, it becomes a factor affecting feedback-eliciting behavior only when the feedback source is another person, or when a person is present to witness the feedback-eliciting behavior (as was the case here). In the latter case, higher levels of internal propensity are correlated with lower levels of total feedback eliciting. In the former case, levels of internal propensity have no effect.

As noted above, state anxiety varied among participants, and was significantly correlated with feedback eliciting (although not in the direction hypothesized). State anxiety was not, however, statistically different across manipulations of social presence and evaluation. Apparently individual differences in anxiety were due to factors other than observer presence or evaluation. It was conjectured that the overall task may have been so salient that it overwhelmed the intended effects of the experimental manipulations. Additional analyses were undertaken to examine this conjecture.

The state anxiety instrument, which was administered immediately following task completion, was designed to measure anxiety at the present moment. It was one of two complementary instruments developed by Spielberger and his colleagues (1983). The other instrument, trait anxiety, was designed to measure "relatively stable individual differences in anxiety-proneness, that is, to differences between people in the tendency to perceive stressful situations as dangerous or threatening..." (Spielberger et al, 1983; p. 1) Participants completed the trait anxiety instrument as part of the survey administered several days or weeks prior to the laboratory experiment. This permitted a comparison between anxiety levels in the classroom, and in the laboratory immediately following the computer task.

A total of 128 subjects completed both the trait anxiety measure (in the classroom) and the state anxiety measure (in the laboratory). When converted to Spielberger's (1983) original metric having a minimum of 20 and a maximum of 80, mean trait anxiety was 42.624 (s.d. = 7.15) and mean state anxiety was 45.394 (s.d. = 10.48); the difference of 2.770 indicated greater anxiety in the laboratory. A paired t-test indicated that the difference was significant ($t = -2.83$, $df = 126$, $p = 0.005$). The Pearson correlation between trait and state anxiety was 0.262 ($p = .003$).

Spielberger (1983) reported the anxiety scores of a normative sample of 855 college students at the University of South Florida. Both the trait anxiety and state anxiety instruments were administered in classroom settings. When weighted by gender, mean trait anxiety for this group was 39.60 (s.d. = 9.80) and state anxiety was 37.89 (s.d. = 11.25). The difference of -1.71 between scores indicates that while the students were completing the state anxiety instrument, they reported feeling less anxiety than they usually or habitually did (i.e., trait anxiety). This is in contrast with the participants in the lab study, who reported higher state anxiety than trait anxiety; that is, the task increased anxiety above the trait, or baseline, level. (Parenthetically, it is worth noting that both trait and state anxiety scores were significantly higher for the Georgia Tech participants than for the USF students. The reasons for the difference are unknown, but could include chronically higher stress levels due to the competitive regimen at Tech, or the characteristics of students who self-select themselves into such a regimen.)

Spielberger (1983) also reported Pearson correlations between trait and state anxiety for normative samples of college students: for males, the correlation was 0.65; for females, 0.59 (significances not reported). For the Tech participants, the correlations were 0.278 for males ($n = 81$, $p = .006$) and 0.121 for females ($n = 47$, $p = .214$). The correlations were transformed to z-scores using Fisher's r-to-z transformation, and the differences between groups were compared using t-tests. The differences were significant between the norm group and the Tech group for both males ($t = -4.29$, $p < .001$) and females ($t = -3.68$, $p < .001$). Compared with the norm group, both the male and female participants in the study had significantly lower correlations between trait anxiety (in the classroom) and state anxiety (in the laboratory).

It has already been shown that the experimental manipulations of social presence and evaluation had no effect. The data just reviewed, however, suggest a main effect for the experiment itself. Compared with their normal levels of anxiety (trait anxiety), participants reported higher levels (state anxiety) after doing the experimental task. In the 855 member norm group, the difference was in the opposite direction; state anxiety was lower than trait anxiety. The participants also displayed lower correlations between trait and state anxiety than the norm group, suggesting the existence of one or more factors affecting the experimental group that accounted for more variance in state anxiety. One of those factors may have been the setup; another, the nature of the task.

Various aspects of the setup, or the sequence of events leading up to doing the task in the lab, may have induced varying levels of anxiety in the participants. The

participants were recruited in their management classes by a middle-aged man who could have been a senior faculty member (but was, in fact, a graduate student). This man heavily emphasized the importance of their participation, both to their school and to managerial science. Each participant reviewed his or her schedule, then selected a one-hour time block on a schedule form and entered his or her name, post office box number, and telephone number. One week before the scheduled session each participant got an individualized reminder notice in his or her post office box. He or she also received a phone call the day before the session. Upon arriving at the lab, each participant was greeted by the same older man who had addressed the class, and was then shown a professionally produced video tape featuring a senior faculty member. In some individuals, this elaborate protocol may have induced the impression that something extremely important was about to happen, and may have induced a level of arousal that overwhelmed the effects of the particular manipulation that he or she was exposed to.

It is also possible that the task itself was so involving and arousing that it overwhelmed the manipulations. Such arousal may have inflated the state anxiety scores; witness scale items such as "I felt jittery" and "I felt at ease." One bit of evidence, although anecdotal, is highly suggestive. While the computer program was being "fine-tuned," a senior feedback researcher on the Georgia Tech faculty was recruited to perform the task, with the specific charge of reviewing the feedback screens and suggesting improvements. Although he was already familiar with the task, he quickly became so engrossed in it that he had to be prompted as follows; "Stop fighting with the task and ask for some feedback!"

In summary: there is evidence that experimental manipulations of observer presence and evaluation produced an indirect effect, via the correlations between the other predictors and feedback eliciting. Further evidence, however, suggests that the experiment as a whole overwhelmed the manipulations.

Having discussed the main effects, to include the failure of the experimental manipulations, we now take up the significant interactions.

The interactions depicted in Figure 10 show that the relationship between state anxiety and feedback eliciting was moderated by level of performance. Overall, those with low anxiety elicited less feedback (either as TFCNT or TFSEC) than those with high anxiety. This is the main effect discussed above. However, poor performance produced more eliciting than good performance, and performance level had a greater effect on those

with high levels of anxiety. The interpretation is straightforward: those who felt anxious while doing the task, and were doing poorly, sought more feedback than those who were less anxious. Because of the different metrics, caution should be used in comparing levels of TFCNT and TFSEC. With that caveat noted, however, the similarity between the two plots in Figure 10 suggests that the interactions reflect a valid relationship among performance, state anxiety, and total feedback eliciting.

The interactions between performance and tolerance for ambiguity as predictors of total feedback eliciting are shown in Figure 11. The appearance of TA as one component of a significant interaction was surprising, since the predictor had displayed neither significant correlations nor significant regression weights with respect to either TFCNT or TFSEC. The nature of the interaction is even more surprising, since it is counterintuitive. Participants who performed poorly, and who described themselves as low in TA, sought less feedback than poor performers who described themselves as high in TA. This much is understandable; those who are comfortable with uncertainty would be expected to engage in less feedback-seeking. The unexpected nature of the interaction appears when considering the participants who performed well. Of these, the ones who described themselves as high in TA did more feedback eliciting than the others. As Figure 11 shows, the phenomenon appears for both operationalizations of total feedback eliciting, TFCNT and TFSEC. As with the interactions between apprehension and performance, such agreement argues for the existence of a valid relationship between the interaction of the two predictors and total feedback-eliciting.

To recapitulate, the interaction shows that (self-reported) high tolerance for ambiguity leads to greater differences in eliciting between high and low levels of performance. One possible interpretation, which must be ventured with diffidence, is that self-reports of TA incorporate an element of one's willingness to elicit information *in the presence of other people*. That is, a high-TA individual may be more comfortable with ambiguity in the abstract, but at the same time is less comfortable with the prospect of engaging in ambiguity-reducing activities that would reveal his or her uncertainty to others. Being unwilling to elicit ambiguity-reducing information in front of other people, high-TA people may be more inclined to elicit it from inanimate sources, such as computers. Further, they may be more likely than low-TA people to do this sort of eliciting when confronted with a need for additional information, such as when their performance is poor. The interaction reported here suggests that this may be the case. Despite their self-reported comfort with ambiguity, high-TAs will take more action to reduce it -- so long as the action

does not involve other people. This is an interesting idea meriting further investigation. An extension of the idea would involve making the salient other person the feedback source, and not merely an observer. In such a case, the interaction between performance and ambiguity tolerance may explain even more variance in feedback eliciting.

The interactions depicted in Figures 12 and 13 are less generalizeable than those previously discussed, because they appear when either TFSEC or TFCNT, but not both, are the dependent variables. The interactions can therefore only be interpreted with respect to certain aspects of feedback eliciting. Figure 12 shows an interaction between internal propensity and task familiarity as predictors of total instances of feedback eliciting. Not surprisingly, the mean level of eliciting is higher for low familiarity than for high; this is the main effect that appeared when TFAM was regressed on the other operationalization of total eliciting, TFSEC (Table 5). As noted in a previous section, participants who described themselves as having a high degree of familiarity with tasks of this type behaved as expected for varying levels of internal propensity; that is, those with lower INT requested feedback less often than those with higher INT. Given a degree of familiarity with the task, those with high INT relied on self-generated, internal feedback in preference to external feedback; those with low INT did the opposite. On the other hand, those who described themselves as being unfamiliar with tasks of this type displayed an opposite effect for levels of INT. Those with higher levels of INT engaged in more feedback eliciting, not less. The result suggests that feedback may have more value for high-INT individuals struggling with an unfamiliar task. Although they prefer to generate their own feedback they do not have the resources to do so; they need information about the task, and about their performance. Unlike low-INTs, these people need feedback not only to improve performance, but also to develop their evaluative abilities vis-a-vis the task. Therefore they do more eliciting - at least initially. As they gain experience, their innate internal feedback propensity regains its importance as a determinant of behavior, and they elicit less feedback than their low-INT counterparts.

Figure 13 shows the interaction between internal propensity and social presence as predictors of the number of seconds spent in feedback eliciting. As noted in the previous section, the participants who performed the task with an observer present behaved in conformity with the hypothesized effect of internal propensity on eliciting; high-INTs engaged in more eliciting, and low-INTs in less. The interesting point of the interaction involves the participants who performed the task in private. For those, the relationship between INT and TFSEC was much weaker, and in the opposite direction. The result

indicates that INT may be social in nature; that is, high-INTs may not prefer to rely on self-generated feedback in general, but only when their feedback-eliciting activities are witnessed by others. If it can be replicated, this result sheds important new light on the nature of internal propensity. INT may be deeply involved with attitudes concerning self-presentation (*vide* the positive correlations with self-esteem and achievement need, Table 4) and self-direction (the correlation with locus of control, Table 4).

Incorporating the interaction terms significantly increased the coefficients of determination (R^2) for both TFCNT and TFSEC. For the first criterion, the increase was 0.162; for the second, 0.182 (Table 5)

The second objective of the research was to see if certain factors affected the relative frequencies of outcome and process feedback eliciting. To recapitulate, the relative frequencies were operationalized in two ways; as the as the ratio of outcome feedback-eliciting instances to total instances (OERCNT), and as the ratio of seconds spent eliciting outcome feedback to the total number of seconds spent eliciting both types of feedback (OERSEC). The results of regressing the hypothesized predictors on the two measures are shown in Table 6. Only one hypothesis (H2) was partially supported by a significant regression weight for performance on OERCNT (Table 6). The coefficient of determination was a 0.164, and there were no significant interactions.

H2 was based on arguments involving impression management. Those doing poorly were expected to ask for less outcome feedback than process feedback because the former was more evaluative in nature, more threatening to one's self-image, and more likely to induce negative evaluations in others. H2 was also motivated by considerations involving the relative usefulness of outcome and process feedback, and in the present context these considerations seem to be more relevant. Participants who were doing very badly -- i.e., those who could not get the screen objects anywhere near the target lines, and had no idea how to do so -- did not need a computer-generated message informing them that their total score was close to zero. They could infer that fact quite accurately from their own observations. What they could not infer, rather, was what they ought to do in order to improve their performance. As a result they tended to ask for process feedback in preference to outcome feedback.

ADDITIONAL ANALYSES AND RESULTS

The results reported above adequately addressed all of the *ad initio* hypotheses. To recapitulate briefly, the results indicated that need for achievement, performance, and state anxiety were the significant main effects. Additional effects of ambiguity tolerance, internal propensity and task familiarity entered via interactions. With respect to feedback seeking (here limited to feedback eliciting), the causal locus seemed to be the task and its effect on participants, rather than the participants' trait-type individual differences. Given these results, was there anything else to be learned from this experiment concerning the causes of feedback eliciting? Fortunately, the answer was yes.

The instrument used to measure individual differences, previously referred to as the survey, also collected three elements of demographic data; sex (or gender), age, and the number of years of full-time work experience. It was plausible that all three had significant effects of feedback-eliciting behaviors. Since it was performed on a computer, the task was technical in nature. It has been suggested that gender-related biases in American education produce lower levels of technical competence in women than in men, which may have led to more diffidence among female participants with respect to both working the task and asking for task feedback. In addition, older participants may have been more proactive with respect to feedback eliciting, due to assertiveness acquired with age. Similarly, those with more work experience may have acquired a greater appreciation for feedback on the job, and may therefore have engaged in more feedback-eliciting in the laboratory. It was therefore decided to examine these three variables as possible predictors, utilizing a purely exploratory approach.

We wished to see if these exploratory predictors would add a significant amount of explained variance to that provided by the predictors examined above. The desire to maximize experimental power resulted in some predictors appearing in the original analyses being dropped. Those deleted had no significant main-effect regression weights with the criteria; external propensity, task-specific internal ability, self-esteem, locus of control and the experimental manipulations SOCIAL and TARGET. Internal propensity was retained because of its theoretically interesting interactions with two different predictors, task familiarity and need for achievement.

The gender variable produced the most interesting findings. The Spearman rank-order correlations between gender and TFCNT ($r_s = -0.39$, $p<.01$) and TFSEC ($r_s = -0.41$,

$p<.01$) indicated that females engaged in less total feedback-eliciting than men (gender was coded female = 1, male = 2). Males reported higher internal propensity ($r_s = 0.17$, $p<.05$) and less anxiety during task execution ($r_s = -.0.23$, $p<.01$). The effect of gender on total feedback-eliciting were supported by multiple regression; regressing TFCNT on the retained variables produced a standardized regression weight of -0.243 (p=.003) for gender, while regressing TFSEC produced a weight of -0.267 (p=.002).

An exploratory investigation of two-way interactions among the predictors produced a new interaction between need for achievement and internal propensity as predictors of total feedback eliciting in seconds, TFSEC (Figure 14). When need for achievement is low, feedback-seeking is essentially the same for subjects with both high and low levels of internal propensity. When need for achievement is high, internal propensity affects feedback-eliciting in the expected manner; those with low internal propensity seek more feedback, those with high internal propensity seek less. This interaction demonstrates an interesting relationship between an established individual difference, achievement need, and a relatively new, feedback-specific individual difference, internal propensity. It appears that achievement need serves as a catalyst for internal propensity; only when a certain level of achievement need is present does internal propensity affect feedback-seeking behavior.

CONCLUSIONS

Overview

Hypothetical antecedents of feedback eliciting (overt feedback seeking) were examined utilizing a computer-based technique that permitted objective measurement of the behavior. Feedback eliciting was operationalized in two complementary ways; as the number of times participants asked for feedback, and as the number of seconds they spent studying it. Two types of feedback were examined; outcome feedback, or information about level of performance, and process feedback, or information about how to improve performance. Separate hypotheses were formulated for each.

The original independent variables, whose inclusion in the study were supported by hypotheses, were as follows: performance, anxiety, external propensity, task-specific internal ability, task familiarity, internal propensity, self-esteem, locus of control, tolerance for ambiguity, and need for achievement. Multivariate regression analysis showed that

only performance, need for achievement, and anxiety were significant as predictors of total eliciting. Task familiarity was a significant predictor of only one operationalization of total eliciting, namely total feedback seconds. Total feedback eliciting, and the ratio of outcome to total feedback eliciting, both decreased as performance level increased. Total eliciting increased at higher levels of need for achievement, and at higher levels of anxiety. Because it was uncorrelated with the experimental treatments, it is possible that anxiety was not related to either social presence or the possibility of being evaluated, but rather to a general level of anxiety induced by the task itself. Comparing levels of trait and state anxiety for the participants in this study with levels in a control group lent weight to this possibility.

State anxiety, tolerance for ambiguity, task familiarity, and social presence were significant as components of two-way interactions with performance and internal propensity. Tolerance for ambiguity interacted with anxiety as a predictor of feedback eliciting, both when eliciting was measured in terms of instances and in terms of total time. Likewise, tolerance for ambiguity interacted with performance as a predictor of both measures of total eliciting. Internal propensity interacted with social presence as a predictor of total eliciting time, and also interacted with task familiarity as a predictor of total eliciting instances.

An exploratory investigation was conducted that involved eliminating some nonsignificant variables and adding age, sex and work experience as predictors. Regressing the revised variable list on the criteria added sex to the list of significant main effects; men engaged in less eliciting than women, both in terms of instances and total time. Regression analysis also identified a significant interaction between achievement need and internal propensity, indicating that internal propensity affected feedback-eliciting only when achievement need was high.

Limitations

Many of the limitations of this study were due to the logistical constraints that fall upon most academic research. Undergraduates at a leading technical university are, arguably, not representative of the general population; although this lack of representativeness should influence effect sizes, and not the nature of the observed relationships among constructs.

A more urgent generalizeability concern involves the nature of the task. The great advantage of using a computerized task was that it permitted measurement of feedback-eliciting behaviors when the participants were completely alone. Any other technique, even

observing from behind a two-way mirror, would have introduced some increment of social presence. The computer also permitted measurement of behavior in terms of instances and in terms of time; the latter would have been extremely labor-intensive for an observer with a stopwatch. The disadvantage involved the special nature of the computer. Many participants performed well, while many others seemed to be completely lost. Some portion of the variance in performance was doubtless due to basic intelligence and manual dexterity, neither of which was controlled. Another portion was probably due to how participants related to computers. As noted above, the "task familiarity" measure related only to participants' familiarity with the graphical user interface. Other factors, such as "computer phobia," were not measured, but these may have been very significant predictors of feedback eliciting. If such factors were present, they almost certainly covaried strongly with performance, which was the strongest single predictor of feedback eliciting.

Suggestions for Further Research

Some of the findings suggest interesting lines for future research. The appearance of need for achievement as a main effect was surprising, especially since it was measured using a five-response questionnaire instead of the Thematic Apperception Test. Further, its internal consistency was the lowest of any scale used (Cronbach's alpha = .606). Much social research is cyclic in nature; ideas come into vogue, go out, and return yet again. These results, limited as they are, suggest that the time has come to revisit achievement need as a significant predictor of human behavior (McClelland, 1961).

The interaction of performance and tolerance for ambiguity as predictors of feedback eliciting (Figure 11) was especially interesting, since those reporting higher ambiguity tolerance were apparently more sensitive to performance. That is, high-TA participants manifested greater differences in eliciting between low and high levels of performance than low-TA participants. In the discussion above it was suggested that ambiguity tolerance may have an element of social aversion; that is, those who say they tolerate ambiguity well may actually mean that they would rather tolerate it than ask somebody else for information. If given an non-human source of feedback, however, they may ask for more information than their low-TA counterparts, and this is what the experimental result suggested. This notion should be investigated, since TA is a venerable and often-used psychometric variable.

The results suggested that internal propensity can affect behavior, but only when disinhibited, or activated, by other variables. INT was never a main effect, but appeared as a component of several interactions. As a predictor of the total time spent eliciting, internal propensity interacted with need for achievement, task familiarity, and observer presence. In each case, the expected effect of internal propensity -- i.e., low eliciting at high INT, and vice versa -- was obtained ONLY at high levels of the other predictor. Viewed as a whole, these results suggest the following course of action. (1) A round of theory development aimed at explicating the role of INT as a moderator should be undertaken. (2) The theory should encompass external propensity and internal ability as well as internal propensity. (3) Future lab studies should focus on these three variables as possible moderators of the main effects identified here. Statistical power should be increased by recruiting more participants, and dropping predictors shown here to be totally ineffectual; e.g., locus of control.

The experimental task was both computer-based, and unfamiliar; a logical extension of this work would involve non-computer-based, familiar tasks. As noted above, some persons have a strong aversion to computers. As further noted, much of the world's work is routine, and observing feedback-seeking behavior in naturalistic settings may be more informative than additional laboratory experimentation. One caution should be entered; regardless of the task or setting, feedback-seeking behaviors should be measured by the researcher using objective criteria, and not uncritically inferred from self-reports. This was the primary concern that motivated the present study. We live in a world of mature information and communication technologies. To the greatest extent possible, within ethical and legal constraints, we should apply these technologies to the study of human behavior, and do so vigorously.

REFERENCES

Ajzen, I. 1989. Attitude structure and behavior. In A.R. Pratkanis, S.J. Breckler and A.G. Greenwald (Eds.), Attitude structure and function: 241-274. Hillsdale, NJ: Lawrence Erlbaum Associates.

Ammons, R.B. 1956. Effects of knowledge of performance: A survey and tentative theoretical formulation. Journal of General Psychology, 69: 279-299.

Chapanis, A. 1964. Knowledge of performance as an incentive in repetitive, monotonous tasks. Journal of Applied Psychology, 48: 263-267.

Annett, J. 1969. Feedback and human behavior. Baltimore: Penguin.

Ashford, S.J. & Cummings, L.L. 1983. Feedback as an individual resource: Personal strategies for creating information. Organizational Behavior and Human Performance, 32: 370-398.

Ashford, S.J. & Cummings, L.L. 1985. Proactive feedback seeking: The instrumental use of the information environment. Journal of Occupational Psychology, 58: 67-79.

Ashford, S.J. & Tsui, A.S. 1991. Self-regulation for managerial effectiveness: The role of active feedback seeking. Academy of Management Journal, 34 (2): 251-280.

Ashford, S.J. 1986. Feedback-seeking in individual adaptation: A resource perspective. Academy of Management Journal, 29: 465-487.

Atkinson, J.W. 1957. Motivational determinants of risk-taking behavior. Psychological Review, 64: 359-372.

Atkinson, J.W. 1964. An introduction to motivation. Princeton, NJ: Van Nostrand.

Battman, W. 1988. Feedback seeking as a means of self-assessment and affect optimization. Motivation and Emotion, 12: 57-86.

Bennett, N., Herold, D.M., & Ashford, S.J. 1990. The effects of tolerance for ambiguity on feedback-seeking behavior. Journal of Occupational Psychology, 63: 343-350.

Boone, C., Brabander, B.D., & Gertis, P. 1991. Perception of control and information search strategy in an investment decisions game. Perceptual and Motor Skills, 72: 391-398.

Bowers, K.S. 1973. Situationism in psychology; An analysis and a critique. Psychological Review, 80: 307-336.

Brockner, J. 1988. Self-esteem at work: Research, theory, and practice. Lexington, MS: Lexington Books.

Budner, S. 1962. Intolerance of ambiguity as a personality variable. Journal of Personality, 30: 29-50.

Campbell, J.D. 1990. Self-esteem and clarity of the self-concept. Journal of Personality and Social Psychology, 59: 538-549.

Cattell, R.B. 1978. The scientific use of factor analysis in behavioral and life sciences. New York: Plenum Press.

Carver, C.S. & Scheier, M.F. 1981. Attention and self-regulation: A control theory approach to human behavior. New York: Springer-Verlag.

Cohen, J. & Cohen, P. 1983. Applied multiple regression/correlation analysis for the behavioral sciences (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cottrell, N.B. 1972. Social facilitation. In C. G. McClintock (Ed.), Experimental social psychology: 185-235. New York: Holt, Rinehard & Winston.

Dipboye, R.L. 1977. A critical review on Korman's self-consistency theory of work motivation and occupational choice. Organizational Behavior and Human Performance, 18: 108-126

DuCette, J. & Wold, S. 1973. Cognitive and motivational correlates of generalized expectancies for control. Journal of Personality and Social Psychology, 26: 420-426.

Earley, P.C., Northcraft, G.B., Lee, C., & Lituchy, T.R. 1990. Impact of process and outcome feedback on the relation of goal setting to task performance. Academy of Management Journal, 33: 87-105.

Feather, N.T. & Volkmer, R.E. 1988. Preference for situations involving effort, time pressure, and feedback in relation to type A behavior, locus of control, and test anxiety. Journal of Personality and Social Psychology, 35: 266-271.

Fedor, D. B. 1991. Recipient responses to performance feedback: A proposed model and its implications. In G. Ferris & K. Rowland (Eds.), Research in personnel and human resources management, Vol 9: 73-120. Greenwich, CT: JAI Press.

Fedor, D.B., Mathieson, K., & Adams, S.M. 1990. The effects of individual and situational factors on different types of feedback seeking behaviors. Paper presented at the annual meeting of the Academy of Management Feedback Symposium, San Francisco.

Fedor, D.B., Rensvold, R.B., & Adams, S.M. 1992. An investigation of factors expected to affect feedback seeking: A longitudinal field study. Personnel Psychology , 45: 779-805.

Feldman, J.M. 1981. Beyond attribution theory: Cognitive processes in performance appraisal. Journal of Applied Psychology, 66 (2): 127-148.

Fishbein, M. & Ajzen, I. 1975. Belief, attitude, intentions, and behavior: An introduction to theory and research. Reading, MA: Addison-Wesley.

Geen, R.G. 1991. Social motivation. In Mark R. Rosenzweig and Lyman W. Porter (Eds.): Annual review of psychology, 42: 377-99. Palo Alto, CA: Annual Reviews Inc.

Glazer, A.N. 1982. Drive theory and social facilitation: A critical reappraisal. British Journal of Social Psychology, 21: 265-282.

Greller, M.M. & Herold, D.M. 1975. Sources of feedback: A preliminary investigation. Organizational Behavior and Human Performance, 13: 244-256.

Hackman, J.R. & Oldham, G.R. 1976. Motivation through the design of work: A test of a theory. Organizational Behavior and Human Performance, 16: 250-279.

Hackworth, D.H. 1989. About face: The odyssey of an American warrior. New York: Simon and Schuster.

Hall, D.T. & Nougaim, E.K. 1968. An examination of Maslow's need hierarchy in an organizational setting. Organizational Behavior and Human Performance, 3: 12-35.

Herold, D.M. & Parsons, C.K. 1985. Individual differences in seeking and processing performance feedback information. Paper presented at the meeting of the American Psychological Association, Los Angeles.

Herold, D.M., Parsons, C.K., & Rensvold, R.B. 1991. The development and role of self-generated feedback in a training environment. Paper presented at the Meetings of the American Psychological Society, Washington, D.C.

Herold, D.M., Parsons, C.K., and Rensvold, R.B. 1994. Individual differences in the generation and processing of performance feedback. Technical Report, Army Research Institute.

Herzberg, F., Mausner, B. & Synderman, B. 1959. The motivation to work. New York: Wiley.

House, R., & Wigdor, L. 1967. Herzberg's dual-factor theory of job satisfaction and motivation. Personnel Psychology, 20: 369-389.

Ilgen, D.R. & Feldman, J.M. 1983. Performance appraisal: A process focus. In L.L. Cummings and B.M. Staw (Eds.), Research in Organizational Behavior (Vol.5, pp. 141-197). Greenwich, CN: JAI Press.

Ilgen, D.R., Fisher, C.D., & Taylor, S.M. 1979. Consequences of individual feedback on behavior in organizations. Journal of Applied Psychology, 64 (4): 349-371.

Ilgen, D.R. & Moore, C. 1982. Combined effects of goal setting and feedback. Technical report 82-1, Purdue University, West Lafayette, ID.

Jablin, F.M. 1979. Supervisor-subordinate communication: The state of the art. Psychological Bulletin, 86: 1201-1222.

Jöreskog, K.G. & Sörbom, D. 1989. LISREL 7: A guide to the program and applications. Chicago: SPSS, Inc.

Kanfer, R., & Ackerman, P.L. 1989. Motivation and cognitive abilities: An integrative /aptitude-treatment interaction approach to skill acquisition. Journal of Applied Psychology, 74: 657-690.

Kim, J.S. 1984. Effect of behavior plus outcome goal setting and feedback on employee satisfaction and performance. Academy of Management Journal, 27: 139-149.

Knight, P.A. & Nadel, J.I. 1986. Humility revisited: Self-esteem, information search, and policy consistency. Organizational Behavior and Human Decision Processes, 38: 196-206

Korman, A.K. 1970. Toward a hypothesis of work behavior. Journal of Applied Psychology, 54: 31-41.

Larson, J.R. 1989. The dynamic interplay between employees' feedback-seeking strategies and supervisors' delivery of performance feedback. Academy of Management Review, 14 (3): 408-422.

Maslow, A. 1943. A theory of human motivation. Psychological Review, 50: 370-396

Matsui, T., Okada, A., & Kakuyama, T. 1982. Influence of achievement need on goal setting, performance, and feedback effectiveness. Journal of Applied Psychology, 67: 645-648.

McClelland, D.C. 1961. The achieving society. Princeton, NJ: Van Nostrand.

McGregor, D. 1960. The human side of enterprise. New York: McGraw-Hill.

McNemar, Q. 1969. Psychological statistics. New York: John Wiley

Meyer, W.U., Folkes, V., & Weiner, B. 1976. The perceived informational value and affective consequences of choice behavior and intermediate difficulty task selection. Journal of Research in Personality, 10: 410-423.

Mischel, W. 1968. Personality and assessment. New York: Wiley

Murray, H.A. 1938. Explorations in personality. New York: Oxford University Press.

Norton, R.W. 1975. Measurement of ambiguity tolerance. Journal of Personality Assessment, 39: 607-619.

Petty, R.E. & Cacioppo, J.T. 1986. The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), Advances in experimental psychology (Vol. 19, pp. 123-205). New York: Academic Press.

Podsakoff, P.M. & Organ, D.W. 1986. Self-reports in organizational research: Problems and prospects. Journal of Management, 12 (4): 531-544.

Rensvold, R.B., Fedor, D.B. and Parsons, C.K. 1993. A test of selected antecedents of perceived feedback-seeking costs. (Working paper, Georgia Institute of Technology).

Rosenberg, M. 1965. Society and the adolescent self-image. Princeton, NJ: Princeton University Press.

Rotter, J.B. 1966. Generalized expectancies of internal versus external control of reinforcement. Psychological Monographs, 80 (1): 1-27.

SPSS Inc. SPSS-X User's Guide, 3rd Edition. Chicago: SPSS Inc.

Siegel, S. 1956. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill.

Spangler, W.D. 1992. Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. Psychological Bulletin, 112; 140-154.

Spielberger, C.D., Gorsuch, R.L., Lushene, R.E., Vagg, P.R., and Jacobs, G.A. 1977. State-Trait Anxiety Inventory-Y. Palo Alto, CA: Consulting Psychologists Press.

Staw, B.M., Sandelands, L.E., & Dutton, J.E. 1981. Threat-rigidity effects in organizational behavior: A multilevel analysis. Administrative Science Quarterly, 26: 501-524.

Steers, R. 1975. Task-goal attributes, N achievement, and supervisory performance. Organizational Behavior and Human Performance, 13: 392-403.

Steers, R.M. & Brownstein, D.N. 1976. A behaviorally based measure of manifest needs in work settings. Journal of Vocational Behavior, 9: 251-266.

Sweeney, P.D. & Wells, L.E. 1990. Reactions to feedback about performance: A test of three competing models. Journal of Applied Social Psychology, 20: 818-834.

Taylor, F.W. 1916. The principles of scientific management. Bulletin of the Taylor Society, September (volume unknown.)

Taylor, S.E., Fisher, C.D., & Ilgen, D.R. 1984. Individuals' reactions to performance feedback in organizations: A control theory perspective. In K.M. Rowland & G.R. Ferris (Eds.), Research in personnel and human resources management (Vol. 2). Greenwich, CN: JAI Press.

Triplett, N.E. 1898. The dynamogenic factors in pacemaking and competition. American Journal of Psychology, 9: 507-533.

Weiner, B. 1978. Achievement strivings. In H. London and J.E. Extner, Jr. (Eds.), Dimensions of personality: 1-36. New York: John Wiley & Sons

Weiner, B., Heckenhausen, H., Meyer, W, & Cook, R.E. 1972. Causal ascriptions and achievement behavior: A conceptual analysis of effort and reanalysis of locus of control. Journal of Personality and Social Psychology, 21: 239-248.

Weiss, H.M. & Knight, P.A. 1980. The utility of humility: Self-esteem, information search, and problem-solving efficiency. Organizational Behavior and Human Performance, 25: 216-223.

Wroten, S.P. 1980. An investigation of the miltidimensionality of the feedback construct within the core dimensions model. Midwest Academy Proceedings.

Table 1
Experimental Manipulations

	Evaluation Target: TASK	Evaluation Target: PARTICIPANT
SOCIAL PRESENCE: Task performed in PRIVATE	**Manipulation I:** Task is a computer-assisted instruction program. Purpose of experiment is to evaluate the task. Task will be performed in private.	**Manipulation II:** Task is a government-developed aptitude test. Purpose of experiment is to evaluate the participant. Results will be correlated with academic grades. Debriefing will include a comparison with other participants.
SOCIAL PRESENCE: Task performed with an OBSERVER	**Manipulation III:** Task is an experiment in computer-assisted instruction. Purpose of experiment is to evaluate the task. Observer will videotape and assist if the computer malfunctions. Videotape will be erased if there are no malfunctions.	**Manipulation IV:** Task is a government-developed aptitude test. Purpose of experiment is to evaluate the participant. Results will be correlated with grades, and compared with those of other participants. Observer will videotape and observe performance. Videotape will be archived and viewed by other researchers. Debriefing will include a comparison with other participants.

Table 2
Scale Reliabilities

SCALES	PARTICIPANTS					
	Survey		Lab		Survey & Lab	
	N	Alpha	N	Alpha	N	Alpha
External Propensity	175	.703	--	--	129	.674
Internal Ability	175	.878	--	--	129	.877
Internal Propensity	175	.617	--	--	129	.641
Self-Esteem	176	.874	--	--	129	.867
Locus of Control	175	.693	--	--	128	.714
Tolerance for Ambiguity	174	.690	--	--	128	.649
Achievement Need	176	.612	--	--	129	.606
State Anxiety	--	--	154	.941	126	.941

Table 3
Pearson and Spearman Correlations (Part 1 of 2)

	SCALES	1	2	3	4	5	6	7	8
1	TFCNT	–	.91**	.36**	.25**	-.66**	.02	-.10	-.10
2	TFSEC	.84**	–	.16*	.05	-.57**	-.03	-.17*	-.21*
3	OERCNT	.36**	.12	–	.90**	-.44**	.03	.01	.02
4	OERSEC	-.00	-.18	.80**	–	-.35**	.01	.01	-.04
5	PERF	-.52**	-.41**	-.41**	-.17	–	.00	.10	.14
6	EXT	.03	-.03	.05	.10	.00	–	.12	.06
7	TIA	-.13	-.13	.01	.08	.04	.13	–	.32**
8	TFAM	-.12	-.20*	.00	.03	.08	.05	.31**	–
9	INT	-.13	-.15	-.02	-.00	.01	-.07	.16	-.09
10	SE	-.03	-.06	.01	.03	-.05	.23**	.32**	.12
11	LOC	-.05	-.06	-.03	.01	.16	.04	.25**	.15
12	TA	-.01	.04	.03	-.04	-.01	-.29**	.08	.12
13	NAch	.11	.12	-.10	-.05	-.01	.34**	.14	.09
14	STANX	.45**	.38**	.16	.01	-.41**	.23**	-.20*	-.12
15	SOCIAL	-.06	.01	-.18	-.14	.01	.05	.07	-.03
16	TARGET	.11	.12	.02	-.06	-.13	.02	-.04	.10

	SCALES	9	10	11	12	13	14	15	16
1	TFCNT	-.15*	-.09	-.05	-.02	.07	.39**	-.08	.10
2	TFSEC	-.12	-.09	-.08	-.04	.04	.39**	-.05	.10
3	OERCNT	.12	.03	-.01	.09	-.10	.11	-.20*	.02
4	OERSEC	.05	.02	-.03	.02	-.13	.11	-.25**	.00
5	PERF	-.01	-.02	.08	.04	-.05	-.48**	.04	-.07
6	EXT	-.09	.18*	.03	-.22**	.33**	.23**	.04	.01
7	TIA	.02	.24**	.19*	.15*	.13	-.19*	.06	-.06
8	TFAM	-.08	.07	.14*	.10	.08	-.13	-.02	.07
9	INT	–	.28**	.06	-.09	.16*	-.13	-.14	-.03
10	SE	.40**	–	.19*	.03	.32**	-.15*	-.06	.11
11	LOC	.09	.28**	–	.19*	.23**	-.22**	-.06	-.15*
12	TA	-.18*	-.07	.14	–	.04	-.27**	.02	.01
13	NAch	.17	.38**	.27**	-.01	–	-.13	.13	.19*
14	STANX	-.14	-.11	-.21*	-.26**	-.09	–	.04	.04
15	SOCIAL	-.09	-.02	-.03	.05	.10	.06	–	.15*
16	TARGET	.00	.15	-.13	.01	.18*	.04	.15	–

* $p < 0.05$, ** $p < 0.01$.

(1) Refer to the List of Abbreviations for scale names.
(2) Pearson correlations below diagonal, Spearman rank-order correlations above diagonal.
(3) Correlations between criteria (scales 1-4) and predictors (scales 5-16) are displayed in boxes.
(4) Significant Spearman correlations whose corresponding Pearson correlations are NOT significant are shown in bold.

Table 4
Regression of Total Feedback Eliciting (TFE)
on Hypothesized Predictors[1]

			Operationalizations of TFE			
			Instances (TFCNT)		Seconds (TFSEC)	
H	PREDICTORS	Relat.[2]	Beta	Sig. T	Beta	Sig. T
1	Performance (PERF)	Neg	**-.366**	**.000**	**-.246**	**.001**
3	External Propensity	Pos	-.090	.314	-.155	.112
4b	Internal Ability (task-spec)	Neg	-.016	.857	-.008	.934
4c	Task Familiarity (TFAM)	Neg	-.070	.388	**-.179**	**.044**
5a	Internal Propensity (INT)	Neg	-.092	.311	-.141	.155
6a	Self-Esteem	Neg	-.032	.743	-.014	.890
7a	Locus of Control	Pos	.045	.593	.042	.648
8	Ambiguity Tolerance (TA)	Neg	.069	.425	.079	.399
9a	Need for Achievement	Pos	**.251**	**.006**	**.259**	**.008**
10a	Social Presence (SOCIAL)	Neg	-.087	.262	-.035	.680
10b	Target of Evaluation	Neg	.002	.979	.046	.598
	State Anxiety (STANX)	Neg	***.351***	***.000***	***.300***	***.004***
			(Equation 1)[3]		(Equation 3)[3]	
	INTERACTIONS					
	PERF X STANX		-1.017	.000	-0.786	.009
	PERF X TA		-1.050	.011	-1.362	.002
	INT X TFAM		-1.901	.025	--	--
	INT X SOCIAL		--	--	-1.451	.019
			(Equation 2)[3]		(Equation 4)[3]	

(1) Results supporting hypotheses shown in **bold**; those contradicting hypotheses shown in ***bold italic***.
(2) Hypothesized relationships with criterion: Neg = negative, Pos = positive.
(3) Parameters of regression equations:

Eq	R^2	Adj. R^2	R^2 chg	df	F	sig. F	F chg	sig F chg.
1	.412	.345	--	12,105	6.132	.0000	--	--
2	.574	.511	.162	15,102	9.153	.0000	12.896	.0000
3	.311	.232	--	12,105	3.949	.0001	--	--
4	.497	.418	.182	15,102	6.603	.0000	12.173	.0000

Table 5

Regression of the Ratio of Outcome to Total Feedback Eliciting (OER) on Hypothesized Predictors[1]

			Operationalizations of OER			
			Instances (OERCNT)		Seconds (OERSEC)	
H	PREDICTORS	Relat.[2]	Beta	Sig. T	Beta	Sig. T
2	Performance	Neg	**-.438**	**.000**	-.209	.058
6b	Self-Esteem	Pos	.016	.872	.046	.675
7b	Locus of Control	Neg	.034	.719	.001	.992
9b	Need for Achievement	Pos	-.104	.290	-.050	.640
11	Evaluation Target	Neg	.027	.769	-.055	.588
			(Equation 1)[3]		(Equation 2)[3]	

(1) Results supporting hypotheses shown in **bold**.
(2) Hypothesized relationships with criterion: Neg = negative, Pos = positive.
(3) Parameters of regression equations:

Eq	R^2	Adj. R^2	df	F	sig. F
1	.217	.164	5,104	4.044	.0006
2	.056	-.009	5,104	.8633	.5382

Table 6

Summary of Support for Hypotheses (Page 1 of 2)

Hypo.	Predictor	Hypothet. Relation to Criterion	Criterion	Criterion Operation-alization	Rank-order Correlation:	Multivariate Regression:	Signifcant interactions with:
1	Performance (PERF)	Negative	TFE	TFCNT	Supports	Supports	EApp, TA
				TFSEC	Supports	Supports	EApp, TA
2	PERF	Negative	OER	OERCNT	Supports	Supports	
				OERSEC	Supports	Not signif.	
3	External Propensity (EXT)	Positive	TFE	TFCNT	Not signif.	Not signif.	
				TFSEC	Not signif.	Not signif.	
4a	Task-specific Internal Ability (TIA)	Positive	TFAM	TFAM	Supports	Not applicable.	
4b	TIA	Negative	TFE	TFCNT	Not signif.	Not signif.	
				TFSEC	Supports	Not signif.	
4c	Task Familiarity (TFAM)	Negative	TFE	TFCNT	Not signif.	Not signif.	
				TFSEC	Supports	Supports	
5a	Internal Propensity (INT)	Negative	TFE	TFCNT	Supports	Not signif.	TFAM
				TFSEC	Not signif.	Not signif.	NAch
5b	TIA X INT		TFE	TFCNT	Not applicable.	Not signif.	
				TFSEC	Not applicable.	Not signif.	
6a	Self-Esteem (SE)	Negative	TFE	TFCNT	Not signif.	Not signif.	
				TFSEC	Not signif.	Not signif.	

Table 6

Summary of Support for Hypotheses (Page 2 of 2)

Hypo.	Predictor	Hypothet. Relation to Criterion	Criterion	Criterion Operation-alization	Rank-order Correlation:	Multivariate Regression:	Signifcant interactions with:
6b	SE	Positive	OER	OERCNT	Not signif.	Not signif.	
				OERSEC	Not signif.	Not signif.	
7a	Locus of Control (LOC)	Positive	TFE	TFCNT	Not signif.	Not signif.	
				TFSEC	Not signif.	Not signif.	
7b	LOC	Negative	OER	OERCNT	Not signif.	Not signif.	
				OERSEC	Not signif.	Not signif.	
8	Tolerance for Ambiguity (TA)	Negative	TFE	TFCNT	Not signif.	Not signif.	PERF
				TFSEC	Not signif.	Not signif.	PERF
9a	Need for Achievement (NAch)	Positive	TFE	TFCNT	Not signif.	**Supports**	
				TFSEC	Not signif.	**Supports**	INT
9b	NAch	Positive	OER	OERCNT	Not signif.	Not signif.	
				OERSEC	Not signif.	Not signif.	
10a	Evaluation Apprehension (EApp)	Negative	TFE	TFCNT	**Contradicts**	**Contradicts**	PERF
				TFSEC	**Contradicts**	**Contradicts**	PERF
10b	EApp	Negative	OER	OERCNT	Not signif.	Not signif.	
				OERSEC	Not signif.	Not signif.	

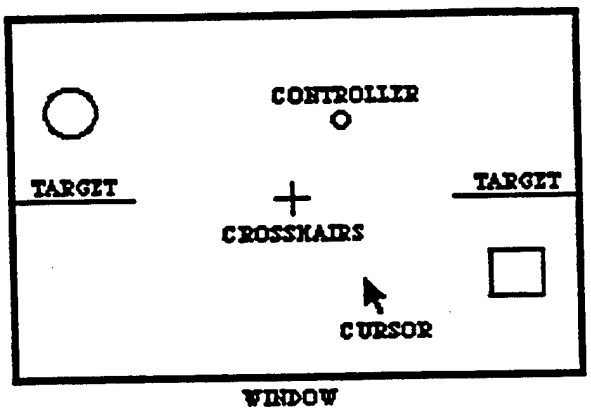


Figure 1
Experimental Task Display

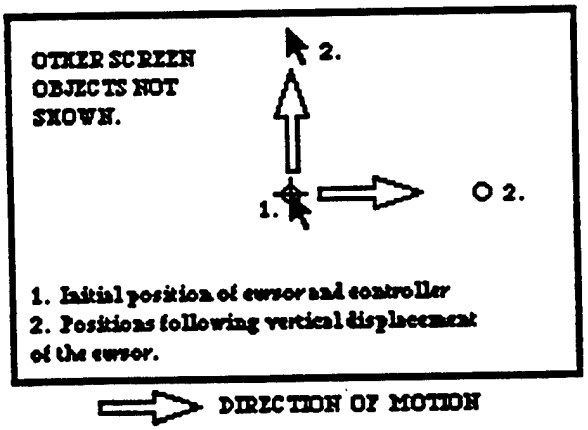


Figure 2
Cursor / Controller Correlation

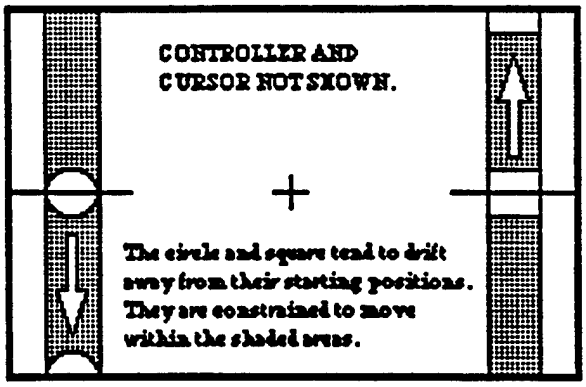


Figure 3
Circle / Square Motion Constraints

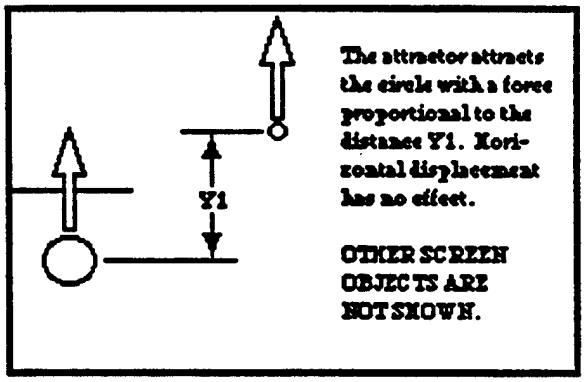


Figure 4
Controlling the Circle

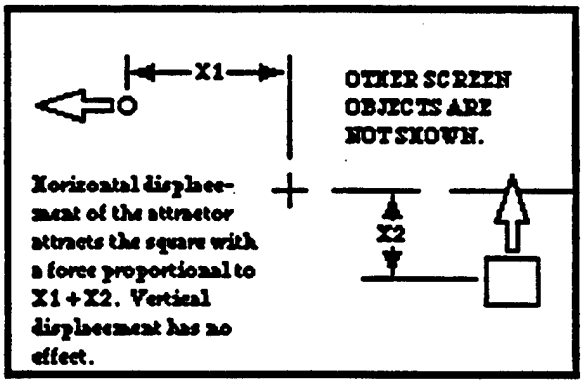


Figure 5
Controlling the Square

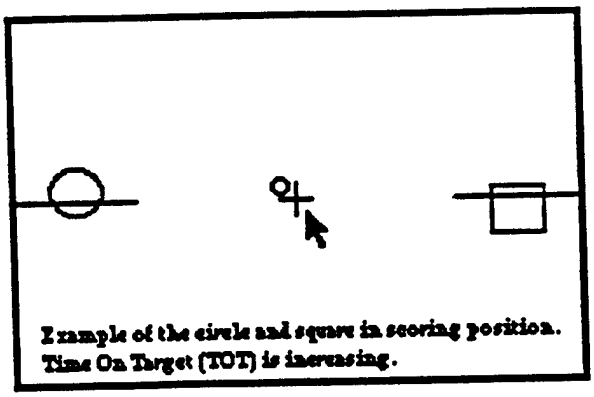


Figure 6
Task Scoring Position

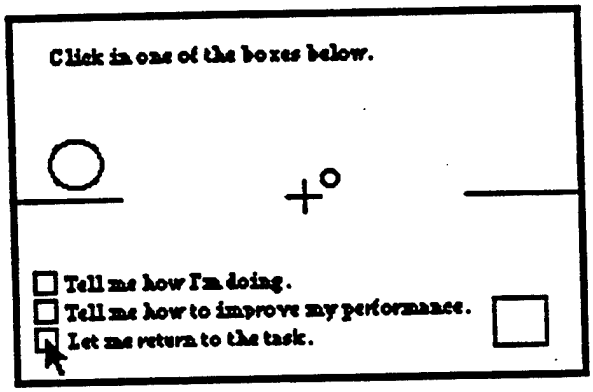
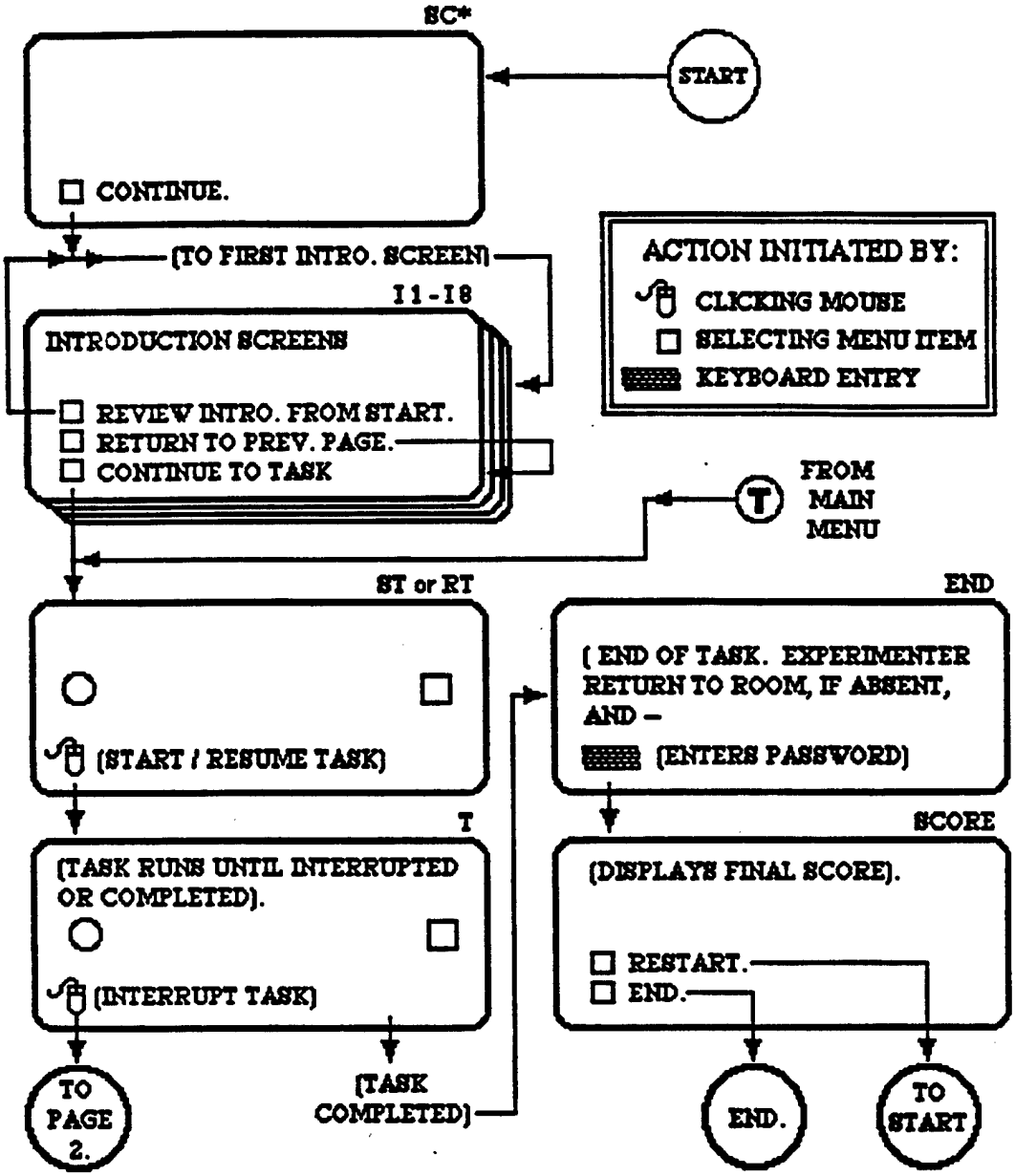


Figure 7
Task Main Menu



*SCREEN NAMES. FOR COMPLETE DESCRIPTIONS SEE APPENDIX A.

Figure 8
Task Flowchart
(Page 1 of 2)

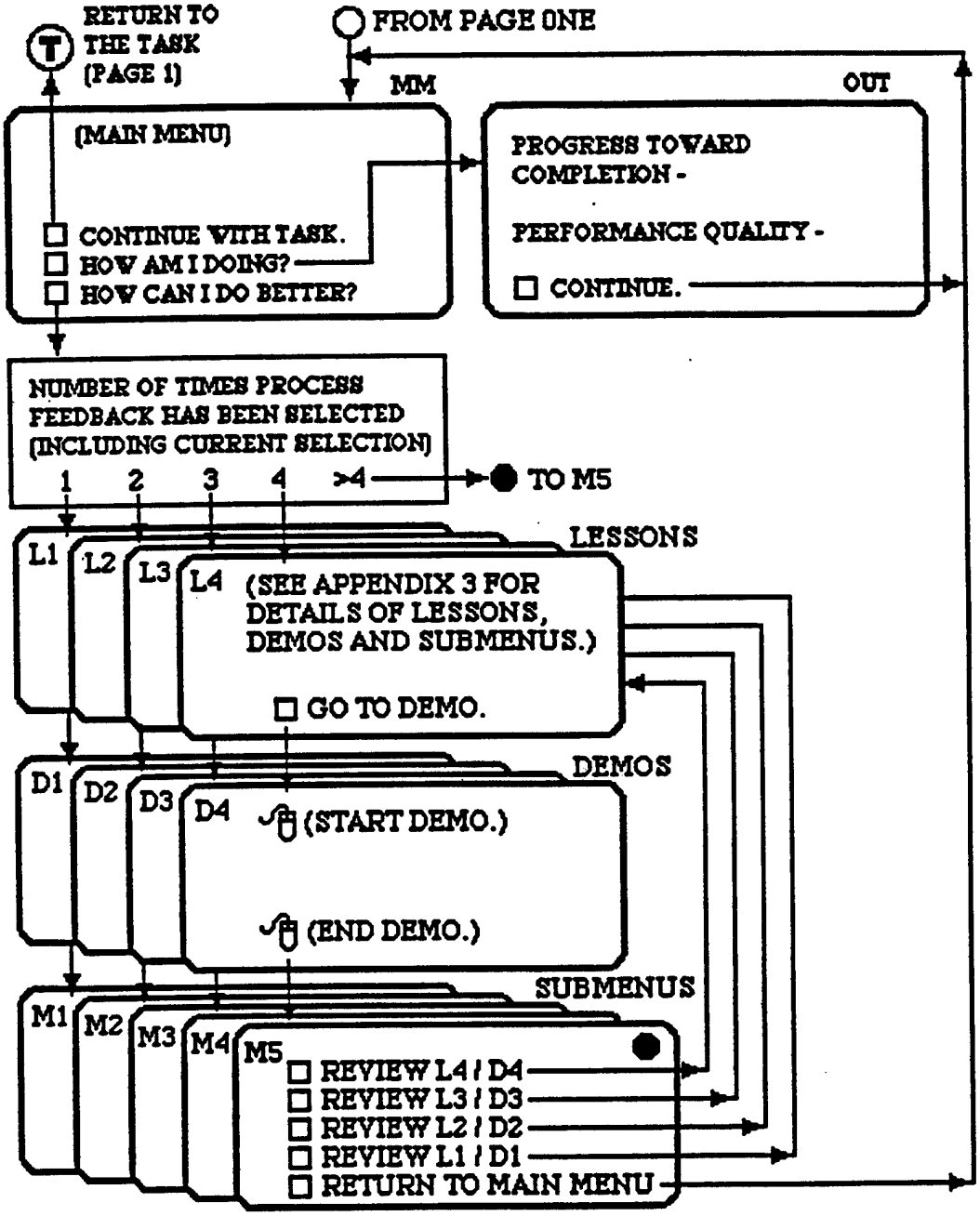
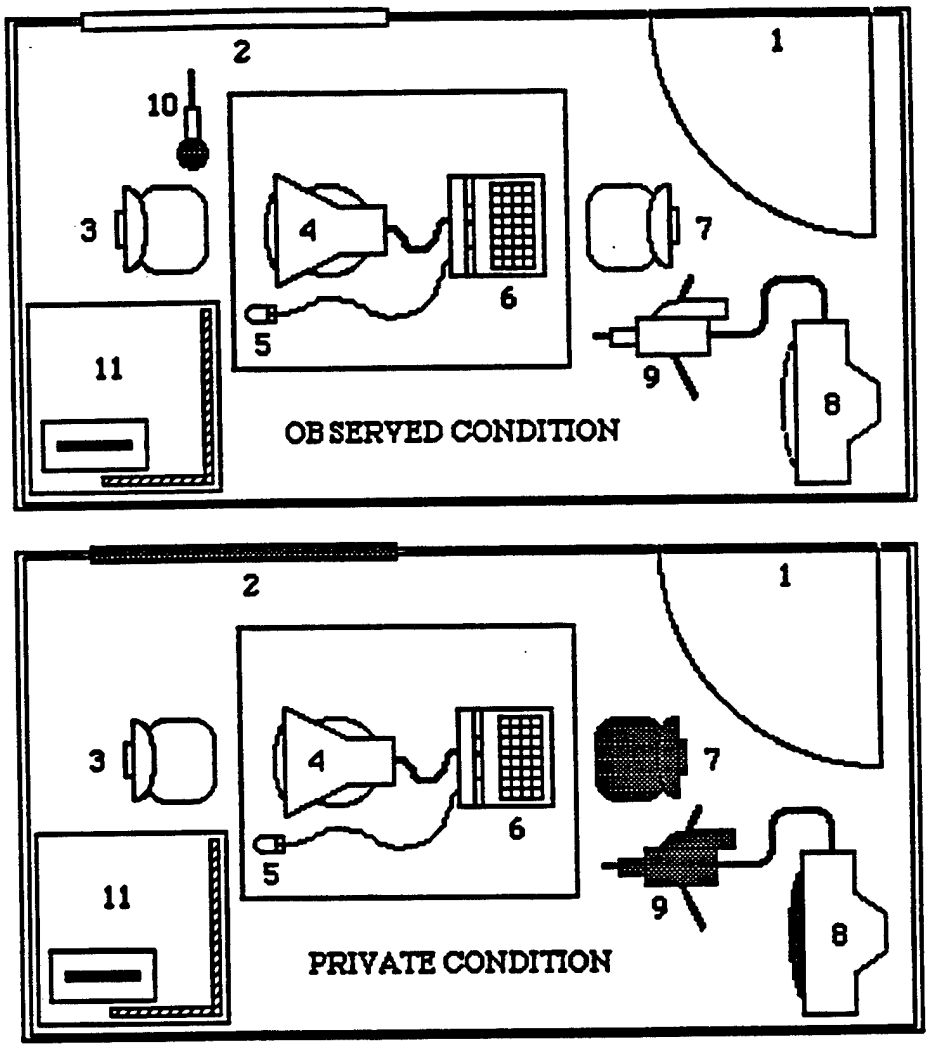


Figure 8
Task Flowchart
(Page 2 or 2)



LEGEND: ☐ In use ■ Not in use

1. Door
2. Two-way mirror
3. Participant's station
4. Participant's monitor
5. Participant's mouse
6. PowerBook computer
7. Experimenter's station
8. TV and VCR on equipment rack
9. Video camera
10. Microphone (suspended)
11. Table with privacy partition (═══) and box for completed questionnaires

Figure 9
Laboratory Setup

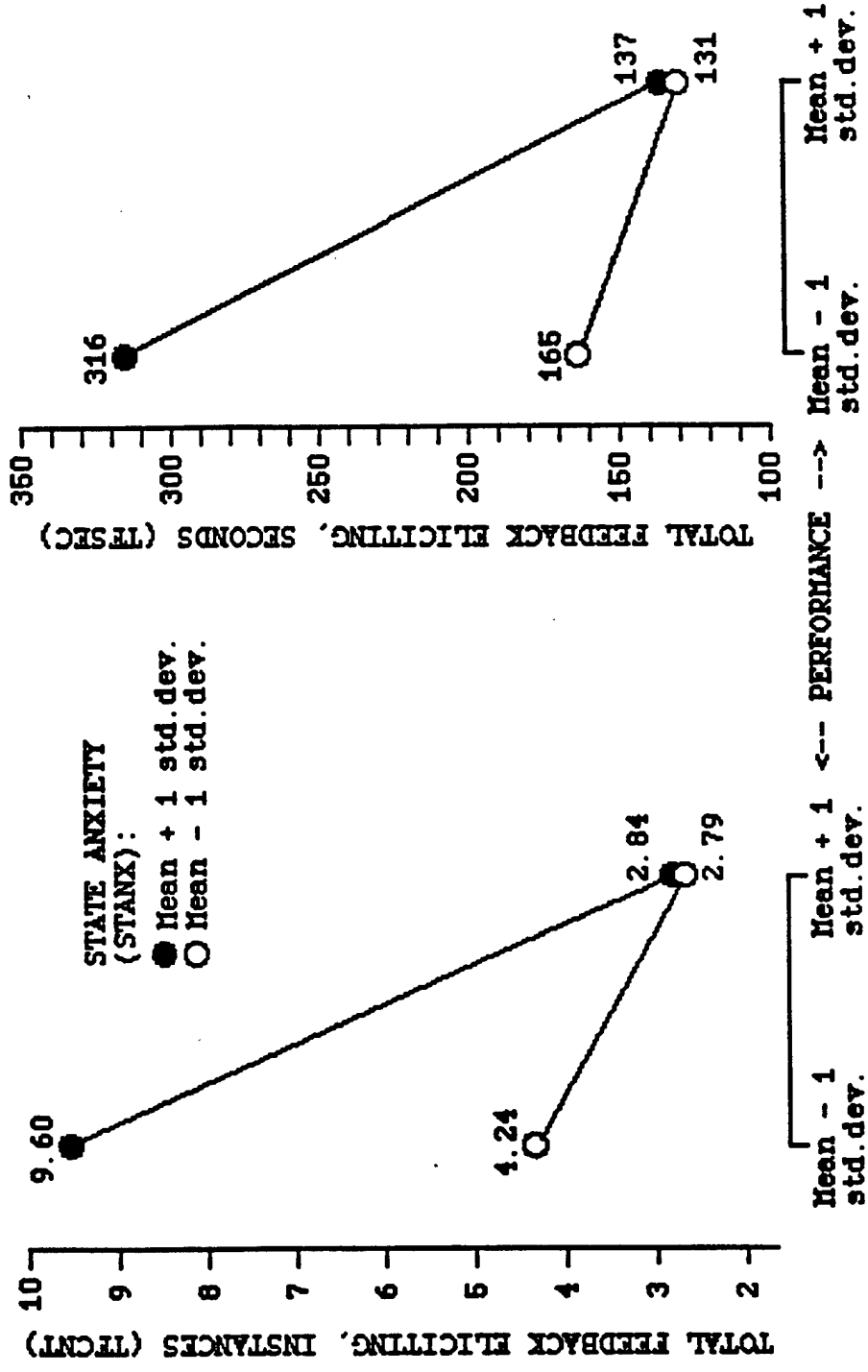


Figure 10.

Interaction of Performance and State Anxiety as Predictors of Total Feedback Eliciting

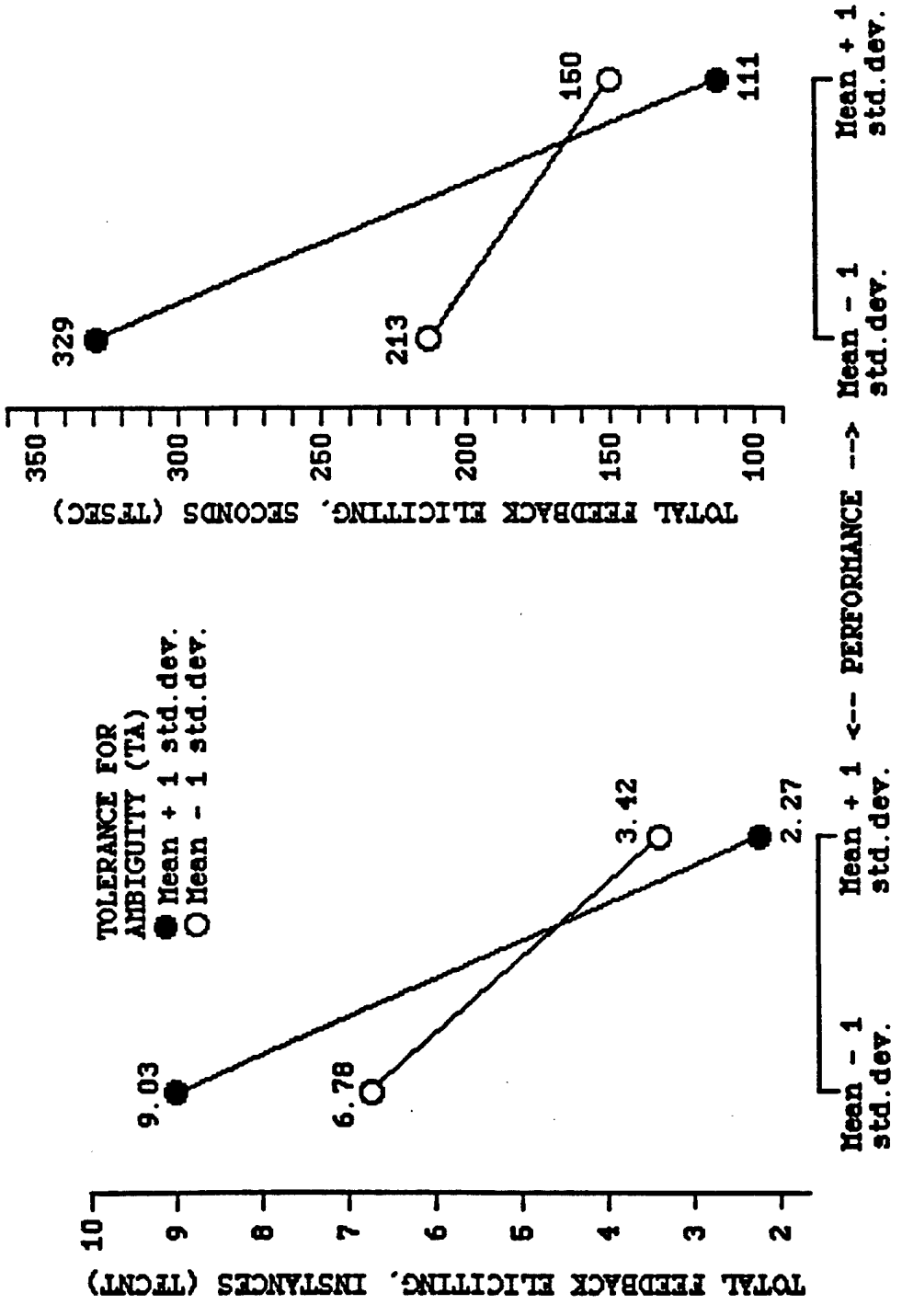


Figure 11.

Interaction of Performance and Tolerance for Ambiguity as Predictors of Total Feedback Eliciting

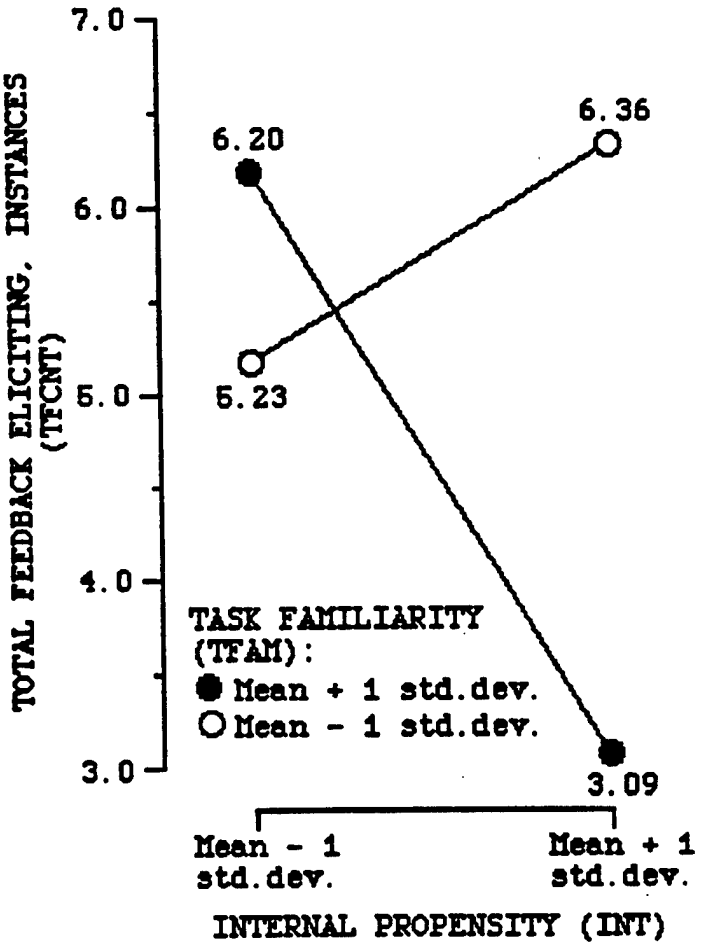


Figure 12.
Interaction of Internal Propensity and Task Familiarity as Predictors of Total Feedback Eliciting (Instances).

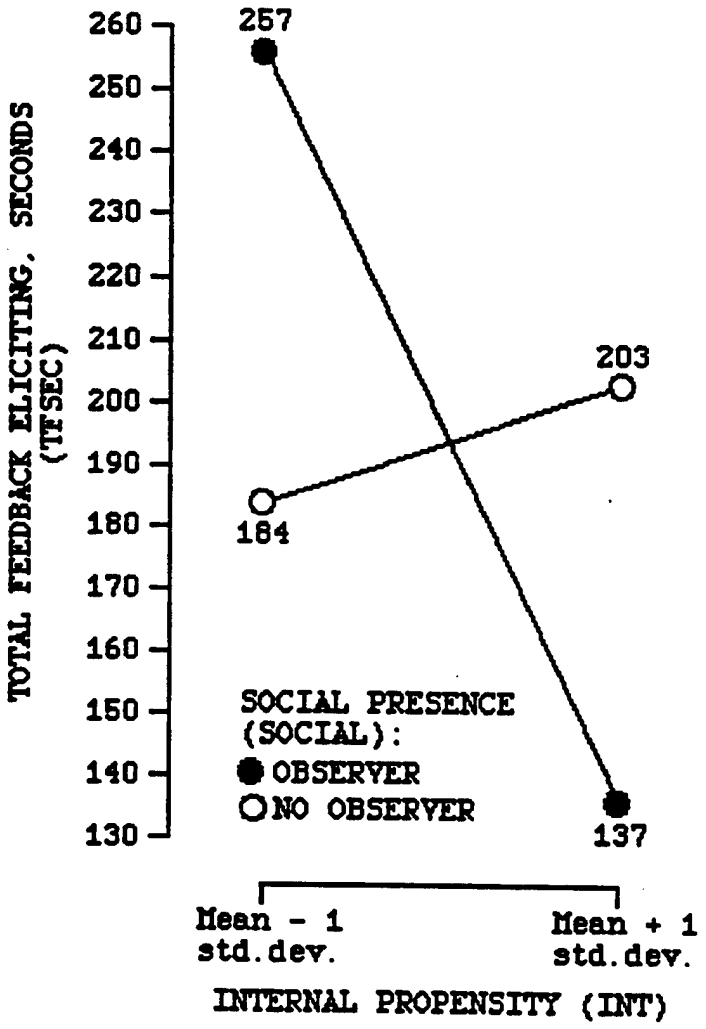


Figure 13.

Interaction of Internal Propensity and Social Presence as Predictors of Total Feedback Eliciting (Seconds)

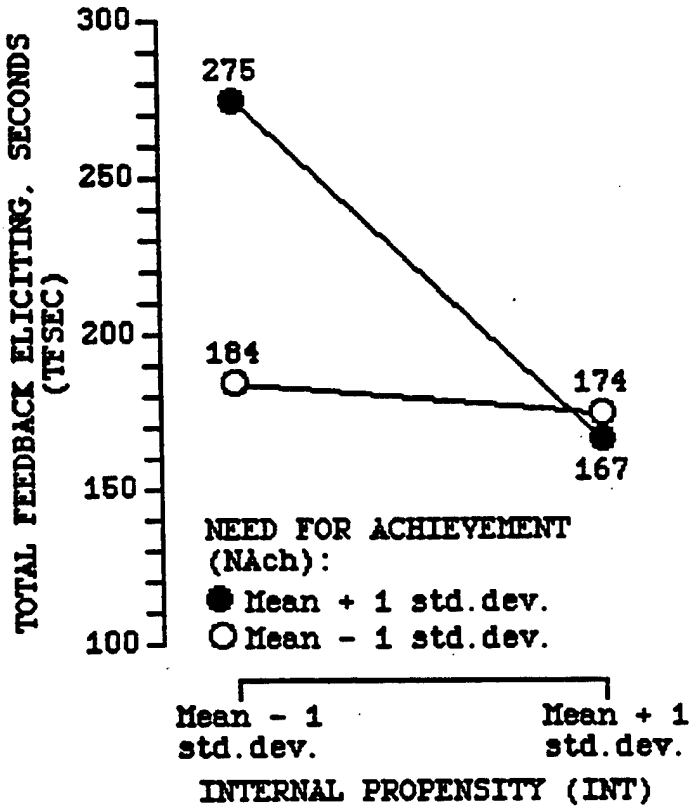


Figure 14.

Interaction of Internal Propensity and Need for Achievement as Predictors of Total Feedback Eliciting (Seconds)

APPENDIX
LIST OF ABBREVIATIONS

COND	The experimental social condition under which the experimental task was performed; either in private (coded 0), or with an observer plus other factors increasing publicness (coded 1). See Figure 9.
DA	Division of attention. The ability to monitor several items simultaneously during task performance.
EXT	External Propensity. A person's preference for, or desire to receive, feedback from an external source.
GUI	Graphical User Interface. A computer interface characterized by pictorial representations of files and applications, which the user selects by using a "pointer" controlled by a "mouse" or similar device. A feature of Macintosh computers and IBM-compatible computers running the Windows operating system .
INT	Internal Propensity. A person's preference for self-generated feedback, as opposed to feedback received from external sources
KURT	Coefficient of Kurtosis.
LOC	Locus of Control
MANIP	Experimental manipulation consisting of the stated purpose of the experimental task; either evaluating the computer program (coded 0) or evaluating the participant, using a computerized aptitude test (coded 1).
MAX	Maximum value of a variable.
MEAN	Mean value of a variable.
MIN	Minimum value of a variable.
N	Number of participants.
NAch	Need for Achievement
OER	Outcome Eliciting Ratio. The ratio of outcome feedback-seeking to total feedback-seeking. Operationalized as Outcome Eliciting Ratio Count (OERCNT) and Outcome Eliciting Ratio Seconds (OERSEC).
OERCNT	Outcome Eliciting Ratio Count. The number of instances the participant halted execution of the experimental task to elicit outcome feedback, divided by the total number of instances he or she elicited both outcome and process feedback. The first operationalization of the outcome eliciting ratio (OER).

OERSEC	Outcome Eliciting Ratio Seconds. The number of seconds the participant spent studying outcome feedback messages presented by the computer, divided by the total number of seconds he or she spent studying both outcome and process feedback messages.
PERF	Performance. Participant's total Time On Target (TOT) divided by the total time spent working on the experimental task.
SE	Self-Esteem.
SKEW	Coefficient of skewness.
STDDEV	Standard Deviation.
TA	Tolerance for Ambiguity
TFAM	Task Familiarity. Operationalized as the participant's self-reported familiarity with computer graphical user interfaces (GUIs), as featured in Macintosh and PC/Windows operating systems.
TFCNT	Total Feedback Count. The number of instances the participant halted execution of the experimental task to elicit either process feedback or outcome feedback from the computer. The first operationalization of total feedback eliciting (TFE).
TFE	Total Feedback Eliciting. The participant's attempts to obtain either outcome feedback or process feedback from the computer. Operationalized as Total Feedback Count (TFCNT) and Total Feedback Seconds (TFSEC)
TFSEC	Total Feedback Seconds. The total time in seconds the participant spent studying the process feedback and outcome feedback messages presented by the computer. The second operationalization of total feedback eliciting (TFE).
TIA	Task-specific Internal Ability. A person's belief in his or her ability to self-generate valid feedback relative to a computer task.
TOT	Time On Target. The number of seconds the participant was able to maintain both the circle and the square on the target lines. See the description of the experimental task.
TT	Total Time. The number of seconds the participant has spent working on the task, in both the on-target and off-target conditions, but excluding time spend viewing introductory screens and obtaining feedback.
WORK	Number of years of full-time work experience, self-reported by experimental participants.